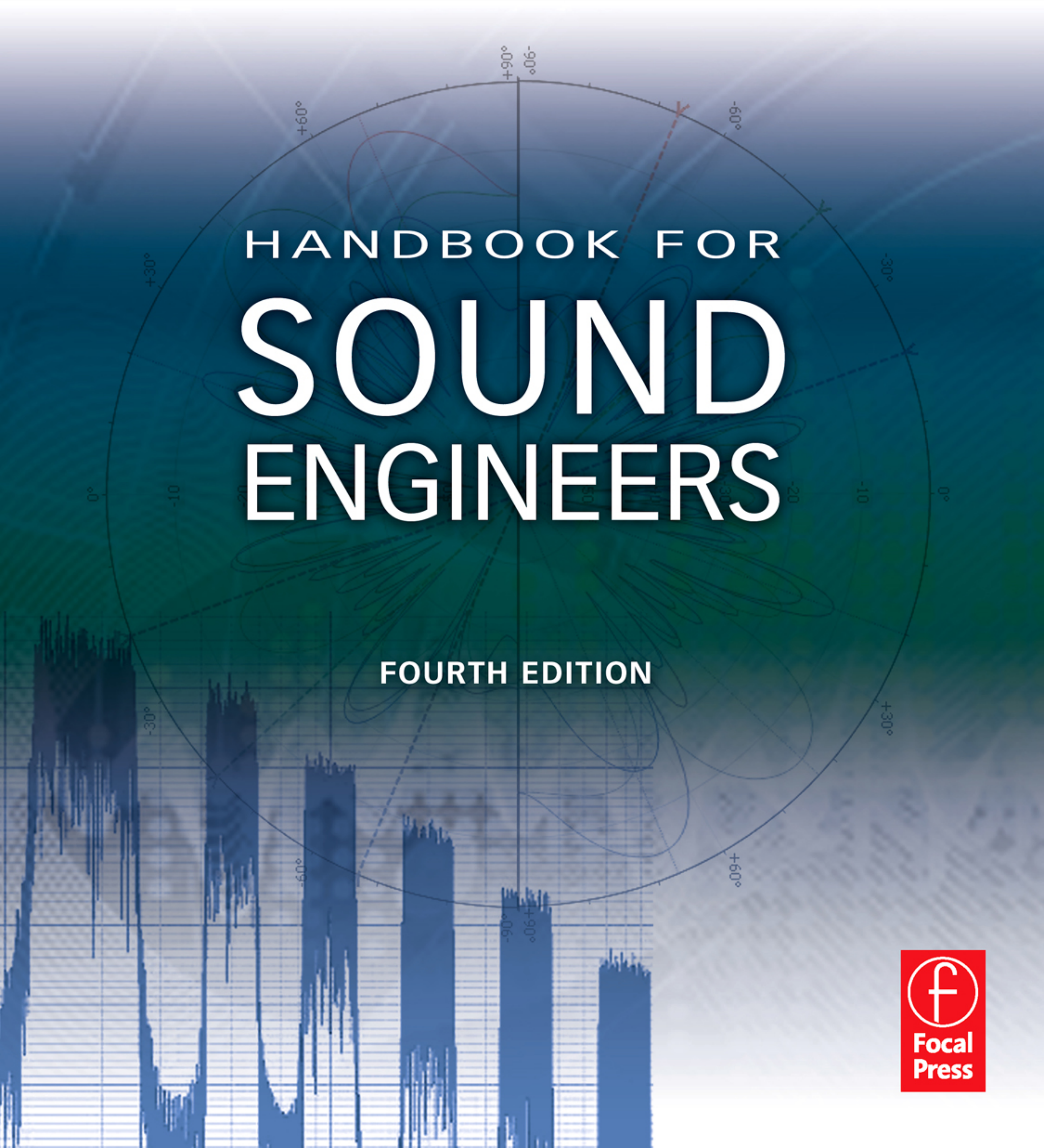
GLEN BALLOU

HANDBOOK FOR SOUND ENGINEERS

FOURTH EDITION



Focal Press

Handbook for Sound Engineers

Fourth Edition

This page intentionally left blank

Handbook for Sound Engineers

Fourth Edition

Glen M. Ballou
Editor

ELSEVIER

Amsterdam • Boston • Heidelberg • London • New York
Oxford • Paris • San Diego • San Francisco • Singapore
Sydney • Tokyo
Focal Press is an imprint of Elsevier

Focal Press

Focal Press is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

Copyright © 2008, Elsevier Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333,E-mail: permissions@elsevier.com. You may also complete your request on-line via the Elsevier homepage (http://elsevier.com), by selecting "Support & Contact" then "Copyright and Permission" and then "Obtaining Permissions."

Library of Congress Cataloging-in-Publication Data Application submitted

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ISBN: 978-0-240-80969-4

For information on all Focal Press publications visit our website at www.elsevierdirect.com

08 09 10 11 5 4 3 2 1

Printed in the United States of America

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Contents

Preface xi
Trademark Acknowledgments xiii
Contributors xv

Part 1—Acoustics

Chapter 1

Audio and Acoustic DNA—Do You Know Your Audio and Acoustic Ancestors? 3
Don and Carolyn Davis

Chapter 2

Fundamentals of Audio and Acoustics. 21
Pat Brown

Chapter 3

Psychoacoustics 41
Peter Xinya Zhang

Chapter 4

Acoustical Noise Control 65
Doug Jones

Chapter 5

Acoustical Treatment for Indoor Areas 95
Jeff Szymanski

Chapter 6

Small Room Acoustics 125
Doug Jones

Chapter 7

Acoustics for Auditoriums and Concert Halls 145
Dr. Wolfgang Ahnert and Hans-Peter Tennhardt

Chapter 8

Stadiums and Outdoor Venues 201
Eugene T. Patronis, Jr.

Chapter 9

Acoustical Modeling and Auralization 213
Dominique J. Chéenne

Part 2—Electronic Components

Chapter 10

Resistors, Capacitors, and Inductors 241
Glen Ballou

Chapter 11

Audio Transformers 273
Bill Whitlock

Chapter 12

Tubes, Discrete Solid State Devices, and Integrated Circuits 309
Glen Ballou, Les Tyler and Wayne Kirkwood

Chapter 13

Heatsinks and Relays 365
Glen Ballou and Henry Villaume

Chapter 14

Transmission Techniques: Wire and Cable 393
Steve Lampen and Glen Ballou

Chapter 15

Transmission Techniques: Fiber Optics 449
Ron Ajemian

Part 3—Electroacoustic Devices

Chapter 16

Microphones 489
Glen Ballou, Joe Ciaudelli and Volker Schmitt

Chapter 17

Loudspeakers 595
Jay Mitchell

Chapter 18

Loudspeaker Cluster Design 645
Ralph Heinz

Part 4—Electronic Audio Circuits and Equipment

Chapter 19

Power Supplies 667
Glen Ballou

Chapter 20

Amplifier Design 701
Eugene T. Patronis, Jr.

Chapter 21

Preamplifiers and Mixers 733
Bill Whitlock, Michael Pettersen and Glen Ballou

Chapter 22

Attenuators 765
Glen Ballou

Chapter 23

Filters and Equalizers 783
Steven McManus

Chapter 24

Delay 805
Steven McManus

Chapter 25

Consoles 817
Steve Dove

Chapter 26

VI Meters and Devices 995
Glen Ballou

Part 5—Recording and Playback

Chapter 27

Analog Disc Playback 1013
George Alexandrovich and Glen Ballou

Chapter 28

Magnetic Recording and Playback 1039
Dale Manquen and Doug Jones

Chapter 29

MIDI 1099
David Mules Huber

Chapter 30

Optical Disc Formats for Audio Reproduction and Recording 1131
Ken Pohlmann

Part 6—Design Applications

Chapter 31

DSP Technology 1159
Dr. Craig Richardson

Chapter 32

Grounding and Interfacing. 1179
Bill Whitlock

Chapter 33

System Gain Structure 1221
Pat Brown

Chapter 34

Sound System Design 1233
Chris Foreman

Chapter 35

Computer Aided Sound System Design 1337
Dr. Wolfgang Ahnert and Stefan Feistel

Chapter 36

Designing for Speech Intelligibility. 1385
Peter Mapp

Chapter 37

Personal Monitor Systems 1413
Gino Sigismondi

Chapter 38

Virtual Systems 1437
Ray Rayburn

Chapter 39

Digital Audio Interfacing and Networking 1457
Ray Rayburn

Chapter 40

Message Repeaters and Evacuation Systems 1513
Glen Ballou and Vic Cappetta

Chapter 41

Interpretation and Tour Group Systems 1529
Glen Ballou

Chapter 42

Assistive Listening Systems 1543
Glen Ballou

Chapter 43

Intercoms 1559
Glen Ballou

Chapter 44

The Fundamentals of Display Technologies 1577
Alan C. Brawn

Chapter 45

Surround Sound 1591
Joe Hull

Part 7—Measurements

Chapter 46

Test and Measurement 1605
Pat Brown

Chapter 47

What's the Ear For? How to Protect It 1631
Les Blomberg and Noland Lewis

Chapter 48

Fundamental and Units of Measurement 1645
Glen Ballou

Index 1689

This page intentionally left blank

Preface

When the first edition of the *Handbook for Sound Engineers* came out in 1987, it was subtitled the new *Audio Cyclopedia* so that people who were familiar with Howard Tremain's *Audio Cyclopedia* would understand that this is an updated version of it. Today, the book stands on its own.

We have seen a tremendous change in the field of sound and acoustics since the first edition of the *Handbook for Sound Engineers* came out. Digital is certainly finding its place in all forms of audio, however, does this mean analog circuitry will soon be a thing of the past? Analog systems will still be around for a long time. After all, sound is analog and the transfer of a sound wave to a microphone signal is analog and from the electronic signal to the sound wave produced by the loudspeaker is analog.

What is changing is our methods of producing, reproducing, and measuring it. New digital circuitry and test equipment has revolutionized the way we produce, reproduce and measure sound.

The *Handbook for Sound Engineers* discusses sound through seven sections, Acoustics, Electronic Components, Electro-Acoustic Devices, Audio Electronic Circuits and Equipment, Recording and Playback, Design Application, and Measurements.

When we listen to sound in different size rooms with different absorptions, reflections, and shape, we hear and feel the sound differently. *The Handbook for Sound Engineers* explains why this occurs and how to control it.

Rooms for speech are designed for intelligibility by controlling shape, reflections and absorption while rooms for music require very different characteristics as blend and reverberation time are more important than speech intelligibility. Multipurpose rooms must be designed to satisfy both speech and music, often by changing the RT60 time acoustically by use of reflecting/absorbing panels or by designing for speech and creating the impression of increased RT60 through ambisonics. Open plan rooms require absorbent ceilings and barriers and often noise masking. Studios and control rooms have a different set of requirements than any of the above.

There are many types of microphones. Each installation requires a knowledge of the type and placement of microphones for sound reinforcement and recording. It is important to know microphone basics, how they work, the various pickup patterns, sensitivity and frequency response for proper installation.

To build, install, and test loudspeakers, we need to know the basics of loudspeaker design and the standard methods of making measurements. Complete systems can be purchased, however, it is imperative the designer understand each individual component and the interrelation between them to design and install custom systems.

With the advent of digital circuitry, sound system electronics is changing. Where once each analog stage decreased the SNR of the system and increased distortion, digital circuitry does not reduce the SNR or increase distortion in the normal way. Digital circuitry is not without its problems however. Sound is analog and to transfer it to a digital signal and change it back to an analog signal does cause distortions. To understand this the *Handbook for Sound Engineers* delves into DSP technology, virtual systems, and digital interfacing and networking.

Analog disk and magnetic recording and playback have changed considerably in the past few years and are still used around the world. The CD has been in the United States since 1984. It is replacing records for music libraries because of its ability to almost instantly locate a spot in a 70+ minute disc. Because a disc can be recorded and rerecorded from almost any personal computer, disc jockeys and home audiophiles are producing their own CDs. Midi is an important part of the recording industry as a standardized digital communications language that allows multiple related devices to communicate with each other whether they be electronic instruments, controllers or computers.

The design of sound systems requires the knowledge of room acoustics, electroacoustic devices and electronic devices. Systems can be single source, multiple source, distributed, signal delayed, installed in good rooms, in bad rooms, in large rooms, or small rooms, all with their own particular design problems. Designing a system which should operate to our specs, but where we did not take into consideration the proper installation techniques such as grounding and common mode signal, can make a good installation poor and far from noise and trouble free. The *Handbook for Sound Engineers* covers these situations, proper installation techniques, and how to design for best speech intelligibility or music reproduction through standard methods and with computer programs.

The new integrated circuits, digital circuitry and computers have given us new sophisticated test gear unthought of a few years ago, allowing us to measure in real time, in a noisy environment, and measure to accuracies never before realized. It is important to know, not only what to measure, but how to measure it and then how to interpret the results.

Fiber optic signal transmission is solidly in the telephone industry and it is becoming more popular in the audio field as a method of transmitting signals with minimum noise, interference and increased security. This does not mean that hard-wired transmission will not be around for a long time. It is important to understand the characteristics of fiber optics, wire and cable and their affects on noise, frequency response and signal loss.

The book also covers message repeaters, interpretation systems, assistive listening systems, intercoms, modeling and auralization, surround sound, and personal monitoring.

The sound level through mega-loudspeakers at rock concerts, through personal iPods, and random noise from machinery, etc. is constantly increasing and damaging our hearing. The *Handbook for Sound Engineers* addresses this problem and shows one method of monitoring noisy environments.

Many of us know little about our audio heritage, therefore a chapter is dedicated to sharing the history of these men who, through their genius, have given us the tools to improve the sound around us.

No one person can be knowledgeable in all the fields of sound and acoustics. This book has been written by those people who are considered, by many, as the most knowledgeable in their field.

Glen Ballou

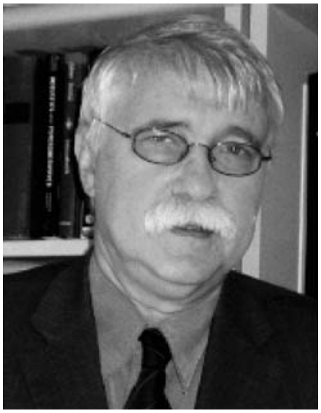
Trademark Acknowledgments

All terms mentioned in this book that are known to be trademarks or service marks are listed below. In addition, terms suspected of being trademarks or service marks have been appropriately capitalized. Focal Press cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Abffusor is a trademark of RPG Diffuser Systems, Inc.
Audio Spotlight is a registered trademark of Holosonic Research Labs, Inc.
AutoCAD is a trademark of Autodesk, Inc.
Beldfoil is a registered trademark of Belden Electronics Division.
Bidiffusor is a trademark of RPG Diffuser Systems, Inc.
Biffusor is a trademark of RPG Diffuser Systems, Inc.
Boogie is a registered trademark of Mesa Engineering.
C-Quam is a registered trademark of Motorola, Inc.
Co-Netic® is a registered trademark of Magnetic Shield Corp.
CobraCAD is a registered trademark of Peak Audio.
CobraNet is a registered trademark of Peak Audio.
dbx is a trademark of DBX Professional Products, AKG, Inc.
Diffractal is a trademark of RPG Diffuser Systems, Inc.
Diffusor is a trademark of RPG Diffuser Systems, Inc.
Dolby is a trademark of Dolby Laboratories Licensing Corporation.
Dolby SR is a trademark of Dolby Laboratories Licensing Corporation.
Electro-Voice, a product brand of Telex Communications, Inc.
Enkasonic is a registered trademark Colbond.
Fiberglas is a registered trademark of Owens Corning.
Flutterfree is a trademark of RPG Diffuser Systems, Inc.
FMX is a registered trademark of Broadcast Technology Partners.
Freon-11 is a trademark of E.I. Dupont de Nemours & Co., Inc.
Holophone® is a registered trademark of Rising Sun ProductionsLimited.
Hot Spot is a trademark of Galaxy Audio.
HyMu® is a registered trademark of Carpenter Technology Corp.
InGenius® is a registered trademark of THAT Corporation.
ISO-MAX® is a registered trademark of Jensen Transformers, Inc.
Isoloop is a registered trademark of 3M Company.
ITE is a trademark of Synergetic Audio Concepts.
Jensen Twin-Servo is a registered trademark of Jensen Transformers, Inc.
Jensen-Hardy Twin-Servo® is a registered trademark of Jensen Transformers, Inc.
Korner-Killer is a trademark of RPG Diffuser Systems, Inc.
LEDE is a trademark of Synergetic Audio Concepts.
Magneplanar is a registered trademark of Magnepan, Inc.
Mumetal® is a registered trademark of Telcon Metals, Ltd.
Mylar is a trademark of E.I. Dupont de Nemours & Co., Inc.
Omniffusor is a trademark of RPG Diffuser Systems, Inc.
OpticalCon® is a registered trademark ofNeutrik AG.
OptiFiber® is a registered trademark of Fluke Networks.
PAR is a trademark of Synergetic Audio Concepts.
PCC is a registered trademark of Crown International, Inc.
Performer is a trademark of Mark of the Unicorn, Inc.
Permalloy® is a registered trademark of B & D Industrial & Mining Services, Inc.
PZM is a registered trademark of Crown International, Inc.
QRD is a trademark of RPG Diffuser Systems, Inc.

QRDX is a trademark of RPG Diffuser Systems, Inc.
Romex® is a registered trademark of General Cable Industries, Inc.
RPG is a trademark of RPG Diffuser Systems, Inc.
SASS is a trademark of Crown International, Inc.
Scotchguard is a trademark of 3M Co.
Series Manifold Technology is a trademark of Electro-Voice, Inc.
Source Independent Measurements is a trademark of Meyer Sound Laboratories, Inc.
Stereo Ambient Sampling System is a trademark of Crown International, Inc.
TEF is a trademark of Gold Line.
Teflon is a trademark of E.I. Dupont de Nemours & Co., Inc.
Tektronix is a registered trademark of Tektronix.
Time Align is a trademark of Ed Long Associates.
Triffusor is a trademark of RPG Diffuser Systems, Inc.
UniCam® is a registered trademark of Corning, Inc.
Variac is a registered trademark of Gen Rad.
Voice Development System is a trademark of Musicode.

Contributors



Professor Dr. Ing. Habil. Wolfgang Ahnert

Dr. Wolfgang Ahnert graduated in Dresden in 1975 and from 1975 to 1990 he worked in an engineering office in Berlin and in 1990 founded ADA Acoustic Design Ahnert. In 2000 he founded SDA Software Design Ahnert GmbH to increase the output in writing source codes for acoustic and electronic applications. In 2002 the ADA foundation was established to support universities and colleges with acoustic software including EASE and the measurement tools EASERA and EASERA SysTune. The Ahnert Feistel Media Group was established in 2006 to coordinate all the above activities.

In 1993 he became an honorary professor at the Hochschule fuer Film und Fernsehen Potsdam, and in 2001 he became an honorary professor at Lomonossov University in Moscow. Since 2005 he has been a visiting professor at the Rensselaer Institute for Architecture in Troy, New York.

Dr. Ahnert has been a member of the AES since 1988 and was made a fellow of the Audio Engineering Society in 1995. He has been a fellow of the Acoustical Society of America since 2005. He is a member of the German DEGA and the British Institute of Acoustics. More than sixty five scientific lectures by Dr. Ahnert have been published.

Dr. Ahnert is one of the authors of the book *Fundamentals of Sound Reinforcement* (1981, and 1984 in Russian) and the book *Sound Reinforcement—Basics and Practice* (1993, and fully updated in English in 2000, in Chinese in 2002, and in Russian in 2003). He wrote with coauthors chapters in the 3rd edition of *Handbook for Sound Engineer*, [*Handbuch der Audiotechnik]* (2008 Springer) and *Akustische Messtechnik* (2008 Springer), both in German.

Ronald G. Ajemian

Ronald G. Ajemian is an instructor at the Institute of Audio Research, New York, NY, where he has been teaching audio technology for over 25 years. Ron was employed with the Switched Services Department of Verizon for 32 years in New York City and recently took an early retirement.

Mr. Ajemian is a graduate of RCA Institute and the school of electrical engineering at Pratt Institute of Brooklyn, NY. He has contributed many articles in the field of audio electronics, telephony, and fiber optics.

Mr. Ajemian is a member of many professional organizations, including the Audio Engineering Society (AES), the Telephone Pioneers of America, Communication Workers of America (CWA), and the Optical Society of America to name a few. He is sometimes referred to as Dr. FO (fiber optics), for his expertise in the field of fiber optics. Mr. Ajemian is also a guest lecturer at NYU and for the Audio Engineering Society. Mr. Ajemian held the position of the Chair of the AES New York Section in 2000–2001 and is currently the Chair of the AES Standards Task Group SC-05-02-F on Fiber Optic Connections.

Mr. Ajemian is owner and consultant of Owl Fiber Optics in New York, specializing in fiber optic technology and education for proaudio/video and broadcast.

George Alexandrovich

George Alexandrovich, born in Yugoslavia, attended schools in Yugoslavia, Hungary, and Germany. After coming to the United States, he studied at the RCA Institute and at Brooklyn Polytech, earning a B.S.E.E. At Telectro Industries Corp., he was involved in the design and development of the first tape recorders and specialized military electronic test and communications equipment.

After service in the Korean war, he ran Sherman Fairchild's private research lab. While with Fairchild Recording Equipment Corp., he designed and manufactured turntables, tonearms, pickups, mixing consoles and amplifiers, equalizers, reverberation chambers, the first light-activated compander, Autoten, Lumiten compressors, limiters, and a line of remote-controlled audio components. He also designed the first professional multichannel portable console, a disk-cutting lathe, and stereo cutters.

As vice president and general manager his responsibilities included designing and manufacturing the Voice of America recording facilities, NBC-TV studio consoles for Johnny Carson, Huntley-Brinkley Newsroom, KNX, KCBS, and other radio stations. When Fairchild Recording Equipment Corp. merged with Robins Corp., George also became involved in manufacturing magnetic tape along with a full line of consumer accessories in the hi-fi market.

At Stanton Magnetics, Inc., as vice president of field engineering and the professional products manager for phonograph cartridge research, he traveled extensively, holding seminars, giving lectures, and conducting conferences.

George was the author of the monthly column "Audio Engineer's Handbook" for *dB Magazine* and of over eighty articles and papers in American and foreign trade journals. He has also presented a number of technical papers at the Audio Engineering Society (AES) conventions and is a fellow and a past governor of the AES. He holds eighteen patents in the audio field and is past chairman of the Electronics Industry Association (EIA) P8.2 Standards Committee.

At the present time George is retired after spending 11 years as principal engineer at ESD/Parker Hannifin Corp. where he conducted R&D in the field of avionics transducers. He is now president and owner of Island Audio Engineering, manufacturing and consulting firm.

Glen Ballou

Glen Ballou graduated from General Motors Institute in 1958 with a bachelor's degree in Industrial Engineering and joined the Plant Engineering Department of the Pratt & Whitney Aircraft Division of United Technologies Corporation. There he designed special circuits for the newly developed tape control machine tools and was responsible for the design, installation, and operation of the 5,000,000 ft^2 plant public address and two-way communication system.

In 1970, Glen transferred to the Technical Presentations and Orientation section of United Technologies' corporate office, where he was responsible for the design and installation of electronics, audio-visual, sound, and acoustics for corporate and division conference rooms and auditoriums. He was also responsible for audio-visual and special effects required for the corporation's exhibit program.

Glen transferred to the Sikorsky Aircraft division of United Technologies as manager of Marketing Communications in 1980, where his responsibilities included the Sikorsky exhibit and special events program, plus operation and design of all conference rooms.

After his retirement from Sikorsky, Glen and his wife, Debra, opened Innovative Communications, a company specializing in sound system design and installation, and technical writing.

Glen is the editor/author of the 1st, 2nd, 3rd, and 4th editions of the *Handbook for Sound Engineers*. He also was a contributing author for *The Electrical Engineering Handbook (*CRC Press). Glen has written many article for Sound and Video Contractor and Church Production magazines.

He has been active in the Audio Engineering Society (AES) as Governor, three times Papers Chairman and four times Facilities Chairman, and Vice Chairman and Chairman of the 1989 AES Convention, for which he received the Governor's award. He was also a member of SMPTE and the IEA.

Les Blomberg

Les Blomberg is a hearing conservation and noise policy expert. He is the inventor of a hearing test to detect temporary changes in hearing ability and is the author of numerous articles and papers on hearing conservation and noise. He is the founding Executive Director of the Noise Pollution Clearinghouse.

Alan C. Brawn CTS, ISF, ISF-C.

Alan Brawn is a principal of Brawn Consulting LLC, an audio-visual consulting, educational development, and market intelligence firm with national exposure to major manufacturers, distributors, and integrators in the industry. He was formerly President of Telanetix and previously National Business Development and Product Marketing Manager, Pro AV Group, Samsung Electronics. Alan is an AV industry veteran with experience spanning over two decades including years managing a commercial AV systems integration company after which he became one of the founding members of Hughes-JVC.

He is a recognized author for leading AV industry magazines and newsletters such as *Systems Contractor News*, *Digital Signage Magazine, Rental & Staging, Display Daily,* and Electrograph's *Emerge.* Brawn is an Imaging Science Foundation Fellow and the Managing Director of ISF Commercial. Alan is CTS certified and an adjunct faculty member of Infocomm sitting on their PETC council. He is an NSCA instructor and Learning Unit Provider. He was the recipient of the Pro AV Hall of Fame recognition from rAVe in 2004.

Alan can be reached at Alan@BrawnConsulting.com.

Pat Brown

Pat Brown is president of Synergetic Audio Concepts, Inc. Syn-Aud-Con conducts educational seminars in audio and acoustics worldwide, as well as private seminars for audio manufacturers and corporations. He has owned and operated several businesses, including design build, retail, and consulting audio firms.

He received an A.A.S. from the University of Louisville in Electrical Engineering Technology. His current duties include teaching, consulting, and authoring technical articles for industry magazines and the Syn-Aud-Con Newsletter, which he and his wife, Brenda, publish from their headquarters in Greenville, Indiana.

Dominique J. Chéenne, Ph.D.

Dominique J. Chéenne holds a Brevet de Technicien Supérieur from the University of Caen, France. He received a Master's degree and a Ph.D., both in Electrical Engineering, from the University of Nebraska, Lincoln. His doctoral dissertation dealt with the modeling of sound propagation over seating surfaces.

In 1979 he founded C & C Consultants, a consulting practice specializing in architectural acoustics and environmental noise control. Since its inception C & C Consultants has provided design services on hundreds of projects nationwide including individual residences, performing arts spaces, schools, offices, factories, churches, as well as local, state, and federal government facilities.

In 1995, Dr. Chéenne accepted an offer to join the faculty of Columbia College. He is currently serving as a tenured member of the faculty in the Audio Arts and Sciences Department where he directs the acoustics program. His main research interests are in the application of computer models to architectural and environmental acoustics issues.

Dr. Chéenne is a member of the Audio Engineering Society and of the Acoustical Society of America.

Don and Carolyn Davis

Don and Carolyn Davis form a unique husband and wife team working in audio, starting in 1951 with The Golden Ear in Lafayette, IN, selling eclectic high-fidelity equipment. Don was Paul Klipsch's President in charge of Vice in the late 50s, and worked for Altec Lansing from 1959 until 1972 where he was co-inventor of 1/3-octave equalization.

The Davises founded Synergetic Audio Concepts in 1973 in response to the growing need in the audio industry for training in the fundamentals of sound reinforcement. Their work in equalization, speech intelligibility, and recording technologies provided the backbone for developing training seminars and workshops. When they turned Syn-Aud-Con over to Pat and Brenda Brown in 1995, Don and Carolyn had been responsible for the education of more than 10,000 sound contractors, designers, and consultants.

Don has authored three books, *Acoustical Tests and Measurements* in 1965; *How to Build Speaker Enclosures* in 1968, co-authored with Alex Badmaieff, which has sold over 200,000 copies, and *Sound System Engineering*, co-authored with his wife, Carolyn, in 1975. They recently completed the 3rd edition with co-author, Dr. Eugene Patronis.

In the process of communicating with the grads of their seminars on audio and acoustics, a quarterly newsletter was established. Most of the newsletter was technical in nature, but it also contained the evolving mindset that bonded teachers and grads into something not originally expected—a fraternity of people dedicated to changing an industry.

After enduring poor sound quality at a meeting of a professional audio society, Don uttered in frustration, *If Bad Sound Were Fatal, Audio Would Be the Leading Cause of Death,* hence the title of Don and Carolyn's book, by that name. They used nontechnical excerpts from the Syn-Aud-Con newsletters, annotated by them in 2003, that go back to the beginning of these changes in the industry up to the Davises' retirement in 1995. The book provides a remarkable insight into entrepreneur teachers' communication with entrepreneur students. The typical attendees at their seminars were already successful in their chosen industry, but knew in their hearts that there was a higher standard and they actively sought it.

They have spent their professional careers writing and lecturing on sound system engineering. The audio industry has generously recognized their efforts as instrumental in the better sound quality we enjoy today. Don and Carolyn are both Fellows of the Audio Engineering Society and have received many awards in the audio industry, including the Distinguished Award in Sound Design and Technology from USITT.

Don and Carolyn reside in Arizona in the winter and on their farm in southern Indiana in the summer.

Steve Dove

Steve Dove is a native to Oxfordshire, England, and is now a resident of Pennsylvania. He has designed mixing consoles in every genre from valves to DSP. On the basis of an early foray with ICs, he became initially a wireman for—until fired for his atrocious soldering—a designer for, and later a director of Alice, a major manufacturer of broadcast, film, and recording consoles. Concurrently, he provided engineering and acoustics expertise to major rock bands, theatres, and studios worldwide.

As a design consultant his clients included Sony Broadcast, Shure Bros., Solid State Logic, Altec Lansing, Clair Bros., Harman/JBL, Crest, and Peavey. He is now Minister of Algorithms for Wheatstone Corporation.

A widely published author with a long list of innovative design techniques and award-winning products to his credit, he is presently realizing an ambition to achieve the class and fluidity of the best of analogue audio in digital signal processing.

Dipl.-Phys. Stefan Feistel

Stefan Feistel studied physics at the University of Rostock and at the Humboldt University of Berlin and graduated with a Master's degree in Theoretical Physics in 2004.

His collaboration with Dr. Wolfgang Ahnert, the owner of ADA Acoustic Design Ahnert, started in 1996. The main target was, and still is, the development of the electro- and room-acoustic simulation software EASE.

In 2000, Wolfgang Ahnert and Stefan Feistel founded the company SDA Software Design Ahnert GmbH, which is dedicated to the development of acoustic modeling and measuring software. In 2003, the nonprofit organization ADA Foundation was established to support universities and colleges with software products such as the simulation software EASE and the measurement tool EASERA. To coordinate all these activities, the Ahnert Feistel Media Group (AFMG) was established in 2006.

Stefan Feistel has authored or co-authored more than thirty articles as a result of the continuously progressing work on software projects like EASE, EASE SpeakerLab, EASE Focus, and EASERA and the related mathematical, numerical, and experimental background studies.

Chris Foreman

Chris Foreman is Vice President and Chief Operating Officer for Community Professional, a loudspeaker manufacturer in Chester, Pennsylvania.

During his pro audio career, Chris has worked in manufacturing, contracting, and tour sound. Chris is widely published, having written numerous magazine articles and technical papers. Chris co-authored the book *Audio Engineering for Sound Reinforcement* with the late John Eargle and he is author of the sound reinforcement chapter of the 1st, 2nd, and 3rd editions of the *Handbook for Sound Engineers* edited by Glen Ballou. He can be reached at chris@proaudioweb.com.

Ralph Heinz

Ralph Heinz joined Renkus-Heinz in 1990 with a background in mechanical engineering. Very quickly he developed a thorough understanding of acoustics and has made many proprietory loudspeaker technologies and patients including Complex-Conic Horns, CoEntrant Topology, and True Array Principle.

David Miles Huber

David Miles Huber has written such books as the industry standard text *Modern Recording Techniques* (www.modrec.com), *The MIDI Manual,* and *Professional Microphone Techniques*. He's also a producer and musician in the dance, chill and downtempo genres, whose CDs, have sold over the million mark. His latest music and collaborations with artists in the United States, United Kingdom, and Europe can be heard at www.51bpm.com and www.MySpace/51bpm.

Joe Hull

Joe Hull started his career in audio in the early 1960s working as a retail salesman in Cambridge, Massachusetts, at a time when some customers still needed convincing that stereo really was better than mono. In 1970, he joined the legendary Advent Corporation, where he worked for 8 years in a variety of sales and marketing positions, concluding that what he liked best was writing about high-technology products, such as the first cassette decks with Dolby B, for a lay audience. He's been doing just that for Dolby Laboratories since 1978, with a few years' hiatus out on his own as a freelance writer. Joe works and lives in San Francisco.

Doug Jones

Doug Jones has worked in recording and small-room acoustics for more than 20 years and still consults on such projects. He has earned a Master's degree from Columbia College, Chicago, where he is Professor of Acoustics and the founding chairman of the Sound Department. There, he directs accredited Bachelor's-degree programs in recording, architectural acoustics, live-sound reinforcement, sound-for-picture, and sound installation.

Mr. Jones is a member of the Acoustical Society of America and the Audio Engineering Society, where he is active in committee work. His publications have appeared in IEEE *Proceedings* and International Computer Music *Proceedings*, and he is an every-month columnist in *Live Sound International* magazine. In addition to his teaching duties at the College, he organizes advanced TEF workshops and other in-service technical training for the audio industry.

Steve Lampen

Steve Lampen has worked for Belden for 17 years and is currently Multimedia Technology Manager. Prior to Belden, Steve had an extensive career in radio broadcast engineering and installation, film production, and electronic distribution. Steve holds an FCC Lifetime General License (formerly a First Class FCC License) and is an SBE Certified Radio Broadcast Engineer. On the data side he is a BICSI Registered Communication Distribution Designer. His latest book, *The Audio-Video Cable Installer's Pocket Guide* is published by McGraw-Hill. His column "Wired for Sound" appears in *Radio World Magazine*.

Noland Lewis

Noland Lewis is a graduate of University of San Francisco and has served as Chief Engineer for HC Electronics/Phonic Ear—a manufacturer of Auditory Trainers and Hearing Aids. Noland has also served on FDA committees relative to hearing aids and has conducted FCC presentations relative to new rule making.

He was the Chief Engineer of HC Electronics (Phonic Ear) where he developed a line of FM Auditory Trainers and Wireless Microphones serving the needs of hearing-impaired children and adults worldwide. He has served on FDA committees relative to hearing aid regulations and submitted new laws that became part of the FCC regulations.

Noland is founder and president of ACO Pacific, Inc., a company he established in 1978—now celebrating 30 years of serving the worldwide community. Over the past 30 years ACO Pacific, Inc. has become a major international supplier of measurement microphones and systems.

Noland created the SLARM™ several years ago to provide a means to assist businesses and communities in the resolution of noise pollution issues.

ACO Pacific is a sustaining member of INCE, INCE USA, AES, ASA, and the Canadian Acoustical Society. He has been a member of numerous committees in these organizations.

Peter Mapp, BSc, MSc, FIOA, FASA, FAES, CPhys, CEng, MinstP, FinstSCE MIEE.

Peter Mapp is principal of Peter Mapp Associates, an acoustic consultancy, based in Colchester, England, that specializes in the fields of room acoustics, electro-acoustics, and sound system design. He holds an honors degree in applied physics and a master's degree in acoustics. He is a Fellow of the Institute of Acoustics, the Acoustical Society of America, and the Audio Engineering Society.

Peter has a special interest in speech intelligibility prediction and measurement and has authored and presented numerous papers and articles on this subject both in Europe and the United States. He currently serves on both British and International Standards committees concerning the sound system design and speech intelligibility.

Peter has been responsible for the design and commissioning of over 500 sound systems, varying from concert halls and theatres, to churches, cathedrals, and other religious buildings, to arenas, stadiums, power stations, and transportation terminals. He is also known for his extensive research work into Distributed Mode Loudspeakers and their application to sound reinforcement.

Peter is a regular contributor to the audio technical press, having written over 100 articles and papers. He is also a contributing author to a number of international reference books including the *Loudspeaker Handbook*. He is currently the chairman of the AES international working group on speech intelligibility and convenor of IEC 60268-16, the standard relating to the measurement of STI and Speech Intelligibility.

Steven McManus

Steven McManus graduated from the University of Edinburgh in 1990 with a B. E. in Electrical and Electronic Engineering. During this time he also worked on a computer-aided navigation and charting project through the University of Dundee.

An early interest in sound recording expanded into a production and installation company that he ran for 8 years. Steve moved to the United States in 1999 after returning to Herriot-Watt University for an update in programming skills.

He was with the TEF division of Gold Line for 6 years working on both software and hardware. He is currently working for Teledyne Benthos in Massachusetts with underwater acoustic communication and navigation systems.

James E. Mitchell

James (Jay) Mitchell began work in the fields of music and audio in 1972. He was awarded a Bachelor of Science in Physics in 1983 and a Master of Science in Physics in 1985, both from the Georgia Institute of Technology.

Since 1985, he has been involved in the design and development of loudspeakers and associated electronics. His design credits include the IMAX Proportional Point Source loudspeaker and the line of loudspeakers currently manufactured by Frazier Loudspeakers.

He is currently president of Frazier Loudspeakers, in Dallas, Texas. He is a member of the Audio Engineering Society.

Eugene T. Patronis, Jr.

Eugene T. Patronis, Jr. is Professor Emeritus of Physics at the Georgia Institute of Technology in Atlanta, Georgia, where he has taught for over 50 years. During the majority of that time he has also served as an industrial and governmental consultant in the fields of acoustics and electronics.

He is the co-author with Don Davis of the 3 edition of *Sound System Engineering.*

Michael Pettersen

Michael Pettersen is the Director of Applications Engineering for Shure Incorporated. Fascinated by sound and audio since building a crystal radio set as a lad, he earned a B.A. in Music Theory from the University of Illinois in 1974.

Employed by Shure since 1976, Michael helped develop and market five different models of automatic microphone mixers. He has presented papers on automatic microphone mixers to the National Association of Broadcasters, the Acoustical Society of America, the National Systems Contractor Association, the European Institute of Acoustics, the Michigan Bar Association, the Voice of America, and others.

He is a member of the Acoustical Society of America and directs Shure's Consultant Liaison Program, a program that informs acoustical consultants worldwide about new product developments from Shure. Michael also plays guitar with jazz big bands in the Chicago area.

Ken C. Pohlmann

Ken Pohlmann serves as an engineering consultant in the design of digital audio systems, in the development of mobile audio systems for automobile manufacturers, and as an expert witness in technology patent litigation. Some of his consulting clients include Alpine Electronics, Analog Devices, AOL Time Warner, Apple, Bertlesmann, Blockbuster, Cirrus Logic, DaimlerChrysler, Ford, Fujitsu Ten, Harman International, Hughes, Hyundai, IBM, Kia, Lexus, Matsushita, Microsoft, Motorola, Nippon Columbia, Onkyo, Philips, Real Networks, Recording Industry Association of America, Samsung, Sony, TDK, Toyota, and United Technologies.

Mr. Pohlmann is a Professor Emeritus at the University of Miami in Coral Gables, Florida, where he served as a tenured full professor and director of the Music Engineering Technology program in the Frost School of Music. He initiated new undergraduate and graduate courses in acoustics and psychoacoustics, digital audio, advanced digital audio, Internet audio, and studio production and founded the first Master's degree in Music Engineering in the United States. Mr. Pohlmann holds B.S. and M.S. degrees in electrical engineering from the University of Illinois in Urbana-Champaign.

Mr. Pohlmann is the author of *Principles of Digital Audio* (5th edition, 2005, McGraw-Hill) and *The Compact Disc Handbook* (2nd edition, 1992, A-R Editions); co-author of *The Master Handbook of Acoustics* (5th edition, 2009, McGraw-Hill), *Writing for New Media* (1997, Wiley), and editor and co-author of *Advanced Digital Audio*

(1991, Sams). He has written over 2,500 articles for audio magazines and is contributing editor, columnist, and blogger for *Sound & Vision* magazine.

Mr. Pohlmann chaired the Audio Engineering Society's International Conference on Digital Audio in Toronto in 1989 and co-chaired the Society's International Conference on Internet Audio in Seattle in 1997. He was presented two AES Board of Governor's Awards (1989 and 1998) and an AES Fellowship Award (1990) by the Audio Engineering Society for his work as an educator and author in the field of audio engineering. In 1991 he was elected to serve on the AES Board of Governors and in 1993 to serve as the AES Vice President of the Eastern U.S. and Canada Region. In 1992 he was awarded the Philip Frost Award for Excellence in Teaching and Scholarship. In 2000 he was elected as a nonboard member of the National Public Radio Distribution/Interconnection Committee. In 2000 he was elected as a member of the Board of Directors of the New World Symphony.

Ray A. Rayburn

Ray A. Rayburn has an ASET from New York Institute of Technology and has worked in the fields of audio, electroacoustics and telecommunications for over 38 years. He has taught audio at Eastman School of Music, Institute of Audio Research, and InfoComm. He designed recording facilities for Broadway Video, Eurosound, Photo-Magnetic Sound, db Studios, and *Saturday Night Live*. His recording credits range from the Philadelphia Orchestra to Frank Zappa. Equipment he has designed includes film dubbers, tape duplicators, specialty telecommunication equipment for the stock brokerage industry, and a polymer analyzer. He has been a consultant on sound systems for the U.S. Senate, U.S. House, Wyoming Senate and House, Georgia Dome, Texas House of Representatives, and University of Hawaii Special Events Arena, among many others.

Currently, Ray is a Senior Consultant with K2 Audio of Boulder, Colorado. He is a member of the Audio Engineering Society (AES), the Acoustical Society of America, the National Systems Contractor's Association, and InfoComm. He is Chairman of the AES Standards Subcommittee on Interconnections and the working group on Audio Connectors. He is a member and former Chairman of the AES Technical Committee on Signal Processing. He was a member of the American National Standards Institute (ANSI) Accredited Standards Committee (ASC) S4 on Audio Engineering. He is a member of the AES Standards working group on Microphone Measurement and Characteristics and of several other Standards working groups. He devotes much of his spare time to educating churches worldwide on how to improve their sound and acoustics through Curt Taipale's Church Sound Check email discussion group.

Ray has a personal web site at www.SoundFirst.com.

Dr. Craig Richardson

Dr. Craig Richardson is a vice president and general manager of Polycom's Installed Voice Business. In this role, Richardson leads the development of installed voice products and recently introduced the Sound Structure audio conferencing products—the first installed conferencing products that provide both mono and stereo echo cancellation capabilities for an immersive conferencing experience.

Prior to joining Polycom, Richardson was president and CEO of ASPI Digital and focused the company on creating the EchoFree™ teleconferencing products for the audio/video integrator marketplace. ASPI's products allowed users to experience full-duplex communication in situations never before thought possible, such as distance learning, telemedicine and courtroom applications. In 2001, ASPI Digital was acquired by Polycom Inc., the leader in unified collaborative communication.

Richardson led algorithm development at ASPI Digital as the director of Algorithm Development working on an advanced low-bit-rate video coder, the MELP military standard voice coder, digital filter design products, computer telephony products, multimedia algorithms for speech, audio, and image processing applications for TMS320 family of digital signal processors. He has written numerous papers and book chapters, has been granted four patents, and is the co-author (with Thomas P. Barnwell, III, and Kambiz Nayebi) of *Speech Coding A Computer Laboratory Textbook*, a title in the Georgia Tech Digital Signal Processing Laboratory Series published by John Wiley & Sons.

Richardson is a member of Tau Beta Pi, Sigma Xi, and a senior member of the IEEE. He was graduated from Brown University with a bachelor's degree in electrical engineering and received both his Master's and doctoral degrees in electrical engineering from the Georgia Institute of Technology.

Gino Sigismondi

Gino Sigismondi, a Chicago native and Shure Associate since 1997, has been active in the music and audio industry for nearly 15 years. Currently managing the Technical Training division, Gino brings his years of practical experience in professional audio to the product training seminars he conducts for Shure customers, dealers, distribution centers, and internal staff. Gino spent over 9 years as a member of Applications Engineering, assisting Shure customers with choosing and using the company's vast array of products, and is the author of the Shure educational publications *Selection and Operation of Personal Monitors*, *Audio Systems Guide for Music Educators*, and *Selection and Operation of Audio Signal Processors*. He was recently awarded status as an Adjunct Instructor by the InfoComm Academy.

Gino earned his B.S. degree in Music Business from Elmhurst College, where he was a member of the jazz band as both guitar player and sound technician. After college, he spent several years working as a live sound engineer for Chicago-area sound companies, night clubs, and several local acts. Gino continues to remain active as a musician and sound engineer, consulting musicians on transitioning to in-ear monitors and expanding his horizons beyond live music to include sound design for modern dance and church sound.

Jeff D. Szymanski

Jeff D. Szymanski is an Acoustical Engineer with Black & Veatch Corporation, a leading global engineering, consulting, and construction company specializing in infrastructure development in energy, water, information, and government market sectors. Jeff's experience covers many areas of acoustics, including architectural acoustics design, industrial noise and vibration control, environmental noise control, and A/V systems design. He has over 13 years of experience in the acoustical manufacturing and consulting industries.

Prior to joining Black & Veatch, Jeff was the Chief Acoustical Engineer for Auralex Acoustics, Inc., where he was the key designer on several award-winning projects. In 2004–5, he and the Auralex team collaborated with members of the Russ Berger Design Group to develop and launch the award-winning pArtScience brand of acoustic treatment products.

Jeff has written and presented extensively on the subjects of acoustics and noise control. He is a full member of the Acoustical Society of America and the Institute of Noise Control Engineering, holds two U.S. patents for acoustic treatment products, is a licensed professional engineer, and he plays guitar pretty well, too.

Hans-Peter Tennhardt

Hans-Peter Tennhardt was born in Annaberg/Erzgebirge, Germany in 1942. From 1962–1968, Mr. Tennhardt studied at the Technical University Dresden, Department of Electrotechnics—Low-current Engineering. His field of study was Electroacoustics with Professor Reichardt.

Mr. Tennhardt graduated in 1968 as a Diploma'd Engineer for Electrotechnics—Low-current Engineering at the TU Dresden, with an extension of the basic study at the Academy of Music Dresden. His Diploma thesis was on the subject "Model Investigations in Town Planning."

From 1968–1991, Mr. Tennhardt was a Scientific Collaborator in the Department Building and Room Acoustics at the Building Academy in Berlin, with Professor Fasold. He became Deputy Head of the Department Building and Room Acoustics of the Institute for Heating, Ventilation, and Fundamentals of Structural Engineering at the Building Academy in 1991.

In 1992 Mr. Tennhardt became Group Leader for room acoustics at the Institute of Building Physics (IBP) of the Frauenhofer Institute Stuttgart, Berlin Branch.

Since then he has been Head of the Department Building Physics and of the Special Section Building and Room Acoustics at the Institute for Maintenance and Modernization of Buildings (IEMB) of the Technical University Berlin.

Bill Whitlock

Bill Whitlock was born in 1944 and was building vacuum-tube electronics at the age of 8 and running a radio repair business at the age of 10. He grew up in Florida, attended St. Petersburg Junior College, and graduated with honors from Pinellas County Technical Institute in 1965. He held various engineering positions with EMR/Schlumberger Telemetry, General Electric Neutron Devices, and RCA Missile Test Project (on a ship in the Pacific) before moving to California in 1971.

His professional audio career began in 1972 when he was interviewed by Deane Jensen and hired as chief engineer by custom console maker Quad-Eight. There he developed Compumix®, a pioneering console automation system, and other innovations. From 1974 to 1981, he designed automated image synthesis and control systems, several theater sound systems, and patented a multichannel PCM audio recording system for producers of the Laserium® laser light show. In 1981, Bill became manager of Electronic Development Engineering for Capitol Records/EMI where he designed high-speed cassette duplicator electronics and other specialized audio equipment. He left Capitol in 1988 to team with colleague Deane Jensen, developing hardware for Spatializer Audio Labs, among others. After Deane's tragic death in 1989, Bill became President and Chief Engineer of Jensen Transformers.

His landmark paper on balanced interfaces was published in the June 1995 *AES Journal*. He is an active member and former chairman of the AES Standards Committee Working Group that produced AES48-2005. Over the years, Bill has presented many tutorial seminars and master classes for the AES as well as presentations to local AES chapters around the world. He suggested major changes to IEC test procedures for CMRR, which the IEC adopted in 2000. He has written numerous magazine articles and columns for *Mix*, *EDN*, *S&VC*, *System Contractor News*, *Live Sound*, *Multi-Media Manufacturer*, and others. Since 1994, he has taught myth-busting seminars on grounding and interfacing to thousands at industry trades shows, Syn-Aud-Con workshops, private companies, and universities—most recently as an invited lecturer at MIT.

Bill is a Fellow of the Audio Engineering Society and a Senior Member of the Institute of Electrical and Electronic Engineers. His patents include a bootstrapped balanced input stage, available as the InGenius® IC from THAT Corporation, and a high-speed, feed-forward AC power regulator, available from ExactPower®. Bill currently designs audio, video, and other signal interfacing devices at Jensen and handles much of their technical support. He also does consulting work as time permits. In his leisure time, Bill enjoys travel, hiking, music, and restoring vintage radio and TV sets.

Peter Xinya Zhang

Peter Xinya Zhang is a faculty member at Columbia College, Chicago. His field is psychoacoustics, specifically sound localization. He received a B.S. in physics from Peking University, P.R. China, and received his Ph.D. in physics from Michigan State University, Lansing.

In his doctoral thesis, he investigated human binaural pitch effects to test various binaural models, and developed a new technique to simulate a 3D sound field for virtual reality with loudspeakers. He has published papers in the *Journal of Acoustical Society of America*, presented at various conferences, and served as co-chair in the session of psychological and physiological acoustics at the 4th Joint Conference of the Acoustical Society of America and Acoustical Society of Japan. He has lectured on psychoacoustics and sound localization at various institutes including Loyola University in Chicago, Peking University (China), Institute of Acoustics at Chinese Academy of Sciences (China), and Shanghai Conservatory (China). Dr. Zhang is a member of the Acoustical Society of America and of the Audio Engineering Society.

This page intentionally left blank

Part 1

Acoustics

This page intentionally left blank

Chapter 1

Audio and Acoustic DNA—Do You Know Your Audio and Acoustic Ancestors?

by Don and Carolyn Davis

Introduction 5
Genesis 5
1893—The Magic Year 6
Bell Laboratories and Western Electric 7
  Harvey Fletcher (1884–1981) 9
  Negative Feedback—1927 10
  Harry Nyquist (1889–1976) 10
  The dB, dBm and the VI 10
  Bell Labs and Talking Motion Pictures. 11
  Motion Pictures—Visual versus Auditory 11
  The Transition from Western Electric to Private Companies 11
Audio Publications 12
The “High” Fidelity Equipment Designers 13
Sound System Equalization 13
Acoustics 14
Professional-Level Audio Equipment Scaled to Home Use 15
  Edwin Armstrong (1890–1954) The Invention of Radio and Fidelity 16
Acoustic Measurements—Richard C. Heyser (1931–1987) 17
Calculators and Computers 18
  The Meaning of Communication 19
Bibliography. 20

This page intentionally left blank

Introduction

This chapter is the DNA of my ancestors, the giants who inspired and influenced my life. If you or a hundred other people wrote this chapter, your ancestors would be different. I hope you find reading the DNA of my ancestors worthwhile and that it will provoke you into learning more about them.

Interest in my audio and acoustic ancestors came about by starting the first independent Hi-Fi shop, The Golden Ear, in Lafayette, Indiana in early 1952. The great men of hi-fi came to our shop to meet with the audio enthusiasts from Purdue: Paul Klipsch, Frank McIntosh, Gordon Gow, H.H. Scott, Saul Marantz, Rudy Bozak, Avery Fisher—manufacturers who exhibited in the Hi-Fi shows at the Hollywood Roosevelt and the Hilton in New York City. We sold our shops in Indianapolis and Lafayette in 1955, and took an extended trip to Europe. In 1958 I went to work for Paul Klipsch as his "President in charge of Vice." Mr. Klipsch introduced me to Lord Kelvin, the Bell Labs West Street personnel, as well as his untrammeled genius.

Altec was the next stop, with my immediate manager being "the man who made the motion picture talk." At Altec I rubbed against and was rubbed against by the greats and those who knew the greats of the inception of the Art. This resulted in our awareness of the rich sense of history we have been a part of and we hope that sharing our remembrance will help you become alert to the richness of your own present era.

In 1972 we were privileged to work with the leaders in our industry who came forward to support the first independent attempt at audio education, Synergetic Audio Concepts (Syn-Aud-Con). These manufacturers represented the best of their era and they shared freely with us and our students without ever trying to "put strings on us."

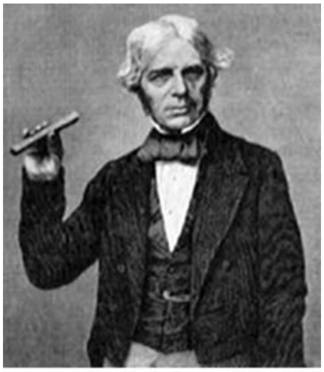
Genesis

The true history of audio consists of ideas, men who envisioned the ideas, and those rare products that represented the highest embodiment of those ideas. The men and women who first articulated new ideas are regarded as discoverers. Buckminster Fuller felt that the terms *realization* and *realizer* were more accurate.

Isaac Newton is credited with "We stand on the shoulders of giants" regarding the advancement of human thought. The word science was first coined in 1836 by Reverend William Hewell, the Master of Trinity College, Cambridge. He felt the term, natural philosopher, was too broad, and that physical science deserved a separate term. The interesting meaning of this word along with entrepreneur-tinkerer allows one a meaningful way to divide the pioneers whose work, stone by stone, built the edifice we call audio and acoustics.

Mathematics, once understood, is the simplest way to fully explore complex ideas but the tinkerer often was the one who found the "idea" first. In my youth I was aware of events such as Edwin Armstrong's construction of the entire FM transmitting and reception system on breadboard circuits. A successful demonstration then occurred followed by detailed mathematical analysis by the same men who earlier had used mathematics to prove its impossibility. In fact, one of the mathematician's papers on the impossibility of FM was directly followed at the same meeting by a working demonstration of an FM broadcast by Armstrong.

The other side of the coin is best illustrated by James Clerk Maxwell (1831–1879), working from the non-mathematical seminal work of Michael Faraday.



Michael Faraday

Michael Faraday had a brilliant mind that worked without the encumbrance of a formal education. His experiments were with an early Volta cell, given him by Volta when he traveled to Italy with Sir Humphry Davy as Davy's assistant. This led to his experiments with the electric field and compasses. Faraday envisioned fields of force around wires where others saw some kind of electric fluid flowing through wires. Faraday was the first to use the terms electrolyte, anode, cathode, and ion. His examination of inductance led to the electric motor. His observations led his good friend, James Clerk Maxwell, to his remarkable equations that defined electromagnetism for all time.

A conversation with William Thomson (later Lord Kelvin) when Thomson was 21 led Faraday to a series of experiments that showed that Thomson's question as to whether light was affected by passing through an electrolyte—it wasn't—led to Faraday's trying to pass polarized light past a powerful magnet to the discover the magneto-optical effect (the Faraday effect). Diamagnetism demonstrated that magnetism was a property of all matter.

Faraday was the perfect example of not knowing mathematics freed him from the prejudices of the day.

James Clerk Maxwell

James Clerk Maxwell was a youthful friend of Faraday and a mathematical genius on a level with Newton. Maxwell took Faraday's theories of electricity and magnetic lines of force into a mathematical formulation. He showed that an oscillating electric charge produces an electromagnetic field. The four partial differential equations were first published in 1873 and have since been thought of as the greatest achievement of the 19th century of physics.

Maxwell's equations are the perfect example of mathematics predicting a phenomenon that was unknown at that time. That two such differing mind-sets as Faraday and Maxwell were close friends bespeaks the largeness of both men.

These equations brought the realization that, because charges can oscillate with any frequency, visible light itself would form only a small part of the entire spectrum of possible electromagnetic radiation. Maxwell's equations predicted transmittable radiation which led Hertz to build apparatus to demonstrate electromagnetic transmission.

J. Willard Gibbs, America's greatest contributor to electromagnetic theory, so impressed Maxwell with his papers on thermodynamics that Maxwell constructed a three-dimensional model of Gibbs's thermodynamic surface and, shortly before his death, sent the model to Gibbs.

G.S. Ohm, Alessandro Volta, Michael Faraday, Joseph Henry, Andre Marie Ampere, and G.R. Kirchhoff grace every circuit analysis done today as resistance in ohms, potential difference in volts, current in amperes, inductance in henrys, and capacity in farads and viewed as a Kirchhoff diagram. Their predecessors and contemporaries such as Joule (work, energy, heat), Charles A. Coulomb (electric charge), Isaac Newton (force), Hertz (frequency), Watt (power), Weber (magnetic flux), Tesla (magnetic flux density), and Siemens (conductance) are immortalized as international S.I. derived units. Lord Kelvin alone has his name inscribed as an S.I. base unit.

As all of this worked its way into the organized thinking of humankind, the most important innovations were the technical societies formed around the time of Newton where ideas could be heard by a large receptive audience. Some of the world's best mathematicians struggled to quantify sound in air, in enclosures, and in all manner of confining pathways. Since the time of Euler (1707–1783), Lagrange (1736–1813), and d'Alembert (1717–1783), mathematical tools existed to analyze wave motion and develop field theory.

Oliver Heavyside

By the birth of the 20th century, workers in the telephone industry comprised the most talented mathematicians and experimenters. Oliver Heaviside's operational calculus had been superseded by Laplace transforms at MIT (giving them an enviable technical lead in education).

1893—The Magic Year

At the April 18, 1893 meeting of the American Institute of Electrical Engineers in New York City, Arthur Edwin Kennelly (1861–1939) gave a paper entitled "Impedance."

Edwin W. Rice

Charles Proteus Steinmetz

That same year General Electric, at the insistence of Edwin W. Rice, bought Rudolph Eickemeyer's company for his transformer patents. The genius Charles Proteus Steinmetz (1865–1923) worked for Eickemeyer. In the saga of great ideas, I have always been as intrigued by the managers of great men as much as the great men themselves. E.W. Rice of General Electric personified true leadership when he looked past the misshapened dwarf that was Steinmetz to the mind present in the man. General Electric's engineering preeminence proceeded directly from Rice's extraordinary hiring of Steinmetz.

Dr. Michael I. Pupin

Dr. Michael I. Pupin of Columbia University was present at the Kennelly paper. Pupin mentioned Oliver Heaviside's use of the word impedance in 1887. This meeting established the correct definition of the word and established its use within the electric industry. Kennelly's paper, along with the ground-work laid by Oliver Heaviside in 1887, was instrumental in introducing the terms being established in the minds of Kennelly's peers.

Arthur Edwin Kennelly

The truly extraordinary Arthur Edwin Kennelly (1861–1939) left school at the age of thirteen and taught himself physics while working as a telegrapher. He is said to "have planned and used his time with great efficiency," which is evidenced by his becoming a member of the faculty at Harvard in 1902 while also holding a joint appointment at MIT from 1913–1924. He was the author of ten books and the co-author of eighteen more, as well as writing more than 350 technical papers.

Edison employed A.E. Kennelly to provide physics and mathematics to Edison's intuition and cut-and-try experimentation. His classic AIEE paper on impedance in 1893 is without parallel. The reflecting ionosphere theory is jointly credited to Kennelly and Heaviside and known as the Kennelly-Heaviside layer. One of Kennelly's Ph.D. students was Vannevar Bush, who ran American's WWII scientific endeavors.

In 1893 Kennelly proposed impedance for what had been called apparent resistance, and Steinmetz suggested reactance to replace inductance speed and wattless resistance. In the 1890 paper, Kennelly proposed the name henry for the unit of inductance. A paper in 1892 that provided solutions for RLC circuits brought out the need for agreement on the names of circuit elements. Steinmetz, in a paper on hysteresis, proposed the term reluctance to replace magnetic resistance. Thus, by the turn of the 20th century the elements were in place for scientific circuit analysis and practical realization in communication systems.

Arthur E. Kennelly's writings on impedance were meaningfully embellished by Charles Proteus Steinmetz's use of complex numbers. Michael Pupin, George A. Campbell, and their fellow engineers developed filter theory so thoroughly as to be worthwhile reading today.

Steinmetz was not at the April 18, 1893 meeting, but sent in a letter of comment which included,

> *It is, however, the first instance here, so far as I know, that the attention is drawn by Mr. Kennelly to the correspondence between the electric term "impedance" and the complex numbers.*
>
> *The importance hereof lies in the following: The analysis of the complex plane is very well worked out, hence by reducing the technical problems to the analysis of complex quantities they are brought within the scope of a known and well understood science.*

The fallout from this seminal paper, its instantaneous acceptance by the other authorities of the day, its coalescing of the earlier work of others, and its utilization by the communication industry within a decade, makes it easily one of the greatest papers on audio ever published, even though Kennelly's purpose was to aid the electric power industry in its transmission of energy.

The generation, transmission, and distribution of electromagnetic energy today has no meaning in itself, but only gains meaning if information is conveyed, thus the tragedy of the use of mankind's precious resources to convey trash.

Nikola Tesla (1856–1943) working with Westinghouse designed the AC generator that was chosen in 1893 to power the Chicago World's Fair

Nikola Tesla

Bell Laboratories and Western Electric

Robert Millikan

Frank Jewett

The University of Chicago, at the end of the turn of the 19th century into the 20th century, had Robert Millikan, America's foremost physicist. Frank Jewett, who had a doctorate in physics from MIT, and now worked for Western Electric, was able to recruit Millikan's top students.

George A. Campbell

George A. Campbell (1870–1954) of the Bell Telephone Laboratories, had by 1899 developed successful "loading coils" capable of extending the range and quality of the, at that time, unamplified telephone circuits. Unfortunately, Professor Michael Pupin had also conceived the idea and beat him to the patent office. Bell Telephone paid Pupin $435,000 for the patent and by 1925 the Campbell-designed loading coils had saved Bell Telephone Co. $100,000,000 in the cost of copper wire alone.

To sense the ability of loading coils to extend the range of unamplified telephone circuits, Bell had reached New York to Denver by their means alone. Until Thomas B. Doolittle evolved a method in 1877 for the manufacture of hard drawn copper, the metal had been unusable for telephony due to its inability to support its own weight over usable distances. Copper wire went from a tensile strength of 28,000 lbs/in^2 with an elongation of 37% to a tensile strength of 65,000 lbs/in^2, an elongation of 1%.

Dr. Thomas Stockham

Campbell's paper in 1922, "Physical Theory of the Electric Wave Filter" is still worthwhile reading today. I remember asking Dr. Thomas Stockham, "Do digital filters ring under transient conditions?" Dr. Stockham, (his wife, Martha, said that she worshipped the air he walked on), replied "Yes" and pointed out that it's the math and not the hardware that determines what filters do. Papers like Campbell's are pertinent to Quantum filters, when they arrive, for the same reasons Dr. Stockham's answer to my question about digital filters was valid.

Bell Telephone Laboratories made an immense step when H.D. Arnold designed the first successful electronic repeater amplifier in 1913.

H.D. Arnold at Bell Laboratories had taken DeForest's vacuum tube, discarded DeForest's totally false understanding of it, and, by establishing a true vacuum, improved materials and a correct electrical analysis of its properties enabled the electronic amplification of voice signals. DeForest is credited with putting a "grid" into a Fleming value.

Sir Ambrose J. Fleming (1848–1945) is the English engineer who invented the two-electrode rectifier which he called the thermionic valve. It later achieved fame as the Fleming valve and was patented in 1904. DeForest used the Fleming valve to place a grid element in between the filament and the plate. DeForest didn't understand how a triode operated, but fortunately Armstrong, Arnold, and Fleming did.

Another Fleming—Sir Arthur (1881–1960)—invented the demountable high power thermionic valves that helped make possible the installation of the first radar stations in Great Britain just before the outbreak of WWII.

The facts are that DeForest never understood what he had done, and this remained true till his death. DeForest was never able, in court or out, to correctly describe how a triode operated. He did however; provide a way for large corporations to challenge in court the patents of men who did know.

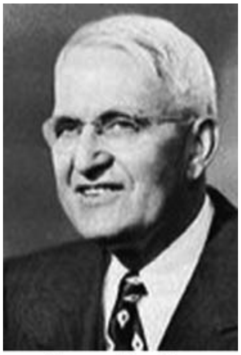
Harold D. Arnold

With the advent of copper wire, loading coils, and Harold D. Arnold's vacuum tube amplifier, transcontinental telephony was established in 1915 using 130,000 telephone poles, 2500 tons of copper wire, and three vacuum tube devices to strengthen the signal.

The Panama Pacific Exposition in San Francisco had originally been planned for 1914 to celebrate the completion of the Panama Canal but the canal was not completed until 1915. Bell provided not only the first transcontinental telephony, but also a public address system at those ceremonies.

The advances in telephony led into recording technologies and by 1926–1928 talking motion pictures. Almost in parallel was the development of radio. J.P. Maxfield, H.C. Harrrison, A.C. Keller, D.G. Blattner were the Western Electric Electrical recording pioneers. Edward Wente's 640A condenser microphone made that component as uniform as the amplifiers, thus insuring speech intelligibility and musical integrity.

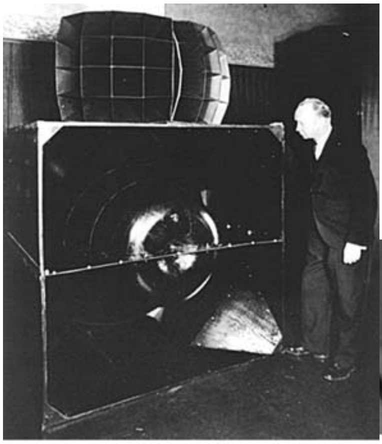
Harvey Fletcher (1884–1981)


Harvey Fletcher

In 1933, Harvey Fletcher, Steinberg and Snow, Wente and Thuras and a host of other Bell Lab engineers gave birth to "Audio Perspective" demonstrations of three-channel stereophonic sound capable of exceeding the dynamic range of the live orchestra. In the late 60s, William Snow was working with John Hilliard at Ling Research, just down the street from Altec. It was a thrill to talk with him. He told me that hearing the orchestra level raised several dB was more astounding to him than the stereophonic part of the demonstration.

Edward Wente

Edward C. Wente and Albert L. Thuras were responsible for full range, low distortion, high-powered sound reproduction using condenser microphones, compression drivers, multicelluar expo nential horns, heavy duty loaded low-frequency enclosures, the bass reflex enclosures, and both amplifiers and transmission lines, built to standards still challenging today. The Fletcher loudspeaker was a three-way unit consisting of an 18 inch low-frequency driver, horn loaded woofer, the incomparable W.E. 555 as a midrange, and the W.E. 597A high-frequency unit.


Albert L. Thuras

W.E. 555 driver

Paul Klipsch with his Cessna 180

Paul Klipsch in the cockpit of his Cessna 180

In 1959, I went with Paul W. Klipsch to Bell Labs where we jointly presented our redo of their 1933 Audio Perspective geometry tests. The demo was held in the Arnold Auditorium and afterward we were shown one of the original Fletcher loudspeakers. Western Electric components like the 555 and 597 are to be found today in Japan where originals sell for up to five figures. It is estimated that 99% of the existing units are in Japan. (As a side note, I genuinely earned a "Distinguished Fear of Flying Cross" with Paul Klipsch in his Cessna 180, the results of which entertained many Syn-Aud-Con classes.

Paul Kilpsch and his assistant in his lab in Hope, Arkansas

The Western Electric 640A was superseded by the 640AA condenser microphone in 1942, still used today as a measurement standard by those fortunate enough to own one. The 640A was a key component in the reproduction of the full orchestra in 1933. When redesigned in 1942 as the 640AA, Bell Labs turned over the manufacturing of the capsule to Bruel and Kjaer as the B&K 4160.

Rice and Kellogg's seminal 1925 paper and Edward Wente's 1925 patent #1,333,744 (done without knowledge of Rice and Kellogg's work) established the basic principle of the direct-radiator loudspeaker with a small coil-driven mass controlled diaphragm in a baffle possessing a broad mid-frequency range of uniform response.

Rice and Kellogg also contributed a more powerful amplifier design and the comment that for reproduced music the level should be that of the original intensity.

Negative Feedback—1927

Harold S. Black

In 1927 Harold S. Black, while watching a Hudson River ferry use reverse propellers to dock, conceived negative feedback for power amplifiers. With associates of the caliber of Harry Nyquist and Hendrik Bode, amplifier gain, phase, and stability, became a mathematical theory of immense use in remarkably diverse technical fields. Black's patent took nine years to issue because the U.S. Navy felt it revealed too much about how they adjusted their big guns and asked that its publication be delayed.

Hendrik Bode

The output signal of an amplifier is fed back and compared with the input signal, developing a "difference signal" if the two signals are not alike. This signal, a measure of the error in amplification, is applied as additional input to correct the functioning of the amplifier, so as to reduce the error signal to zero. When the error signal is reduced to zero, the output corresponds to the input and no distortion has been introduced. Nyquist wrote the mathematics for allowable limits of gain and internal phase shift in negative feedback amplifiers, insuring their stability.

Harry Nyquist (1889–1976)

Harry Nyquist

Harry Nyquist worked at AT&T's Department of Development and Research from 1917 to 1934 and continued when it became Bell Telephone Laboratories in that year, until his retirement in 1954.

The word inspired means "to have been touched by the hand of God." Harry Nyquist's 37 years and 138 U.S. patents while at Bell Telephone Laboratories personifies "inspired." In acoustics the Nyquist plot is by far my favorite for first look at an environment driven by a known source. The men privileged to work with Harry Nyquist in thermal noise, data transmission, and negative feedback all became giants in their own right through that association.

Nyquist worked out the mathematics that allowed amplifier stability to be calculated leaving us the Nyquist plot as one of the most useful audio and acoustic analysis tools ever developed. His cohort, Hendrik Bode, gave us the frequency and phase plots as separate measurements.

Karl Kupfmuller (1897–1977) was a German engineer who paralleled Nyquist's work independently, deriving fundamental results in information transmission and closed-loop modeling, including a stability criterion. Kupfmuller as early as 1928 used block diagrams to represent closed-loop linear circuits. He is believed to be the first to do so. As early as 1924 he had published papers on the dynamic response of linear filters. For those wishing to share the depth of understanding these men achieved, Ernst Guillemin's book, Introductory Circuit Theory, contains clear steps to that goal.

Today's computers as well as digital audio devices were first envisioned in the mid-1800s by Charles Babbage and the mathematics discussed by Lady Lovelace, the only legitimate daughter of Lord Byron. Lady Lovelace even predicted the use of a computer to generate musical tones. Harry Nyquist later defined the necessity for the sampling rate for a digital system to be at least twice that of the highest frequency desired to be reproduced.

Nyquist and Shannon went from Nyquist's paper on the subject to develop "Information Theory." Today's audio still uses and requires Nyquist plotting, Nyquist frequency, the Nyquist-Shannon sampling theorem, the Nyquist stability criterion, and attention to the Johnson-Nyquist noise. In acoustics the Nyquist plot is by far my favorite for first look at an environment driven by a known source.

The dB, dBm and the VI

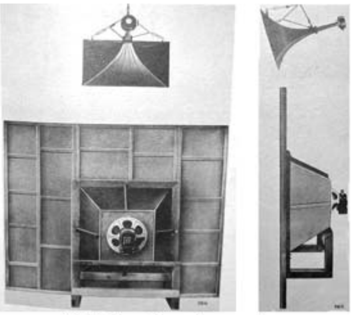
The development of the dB from the mile of standard cable by Bell Labs, their development and sharing of the decibel, dB, the dBm, and the VU via the design of VI devices changed system design into engineering design.

Of note here to this generation, the label VU is just that, VU, and has no other name, just as the instrument is called a volume indicator, or VI. In today's world, a majority of technicians do not understand the dBm and its remarkable usefulness in system design. An engineer must know this parameter to be taken seriously.

Bell Labs and Talking Motion Pictures

Bell Telephone Laboratories by the mid to late 1930s had from the inception of talking motion pictures in 1927–1928 brought forth the condenser microphone, exponential high frequency horns, exponential low frequency loudspeakers, compression drivers, the concepts of gain and loss, the dBm, the VU, in cooperation with the broadcasting industry, and installed sound in 80% of the existing theater market.

Yes, there were earlier dabblers thinking of such ideas but their ideas remained unfulfilled. What generated the explosive growth of motion picture sound—even through the deepest depression—was that only (1) entertainment, (2) tobacco, and (3) alcohol were affordable to the many and solaced their mental depression.


W.E. loudspeaker system installed in the late 1930s

For physicists, motion picture sound was that age's "space race" and little boys followed the sound engineers down the street saying, "He made the movie talk." Dr. Eugene Patronis sent me a picture of the W.E. loudspeaker system installed in the late 1930s in which the engineer had actually aligned the H.F. and L.F. drivers. Dr. Patronis had worked in the projector booth as a teenager. He later designed an outstanding loudspeaker system for the AMC theater chain that was aligned and installed above rather than behind the screen, thereby allowing much brighter images. The system maintained complete spatial location screen-center for the audio.

Motion Pictures—Visual versus Auditory

The first motion pictures were silent. Fortunes were made by actors who could convey visual emotion. When motion pictures acquired sound in 1928, a large number of these well-known personalities failed to make the transition from silent to sound. The faces and figures failed to match the voices the minds of the silent movie viewers had assigned them. Later, when radio became television, almost all the radio talent was able to make a transition because the familiar voices predominated over any mental visual image the radio listener had assigned to that performer.

Often, at the opera, the great voices will not look the part but, just a few notes nullify any negative visual impression for the true lover of opera, whereas appearance will not compensate for a really bad voice.

The Transition from Western Electric to Private Companies

A remarkable number of the giants in the explosion in precision audio products after WWII were alumni of Western Electric-Bell Labs, MIT, and General Radio, and in some cases, all three.

In 1928, a group of Western Electric engineers became the Electrical Research Products, Inc. (ERPI), to service the theaters. Finally a consent decree came down, as a result of litigation with RCA, for W.E. to divest itself of ERPI. At this point the engineers formed All Technical Services or Altec. That is why it is pronounced all-tech, not al-tech. They lived like kings in a depressed economy. As one of these pioneer engineers told me, "Those days were the equivalent of one ohm across Fort Knox." They bought the W.E. Theater inventory for pennies on the dollar.

John Hilliard

The motion picture company MGM had assembled, via Douglas Shearer, head of the sound department, John Hilliard, Dr. John Blackburn, along with Jim Lansing, a machinist, and Robert Stephens, a draftsman. A proprietary theater loudspeaker was named the Shearer horn. Dr. Blackburn and Jim Lansing did the high frequency units with Stephens, adapting the W.E. multicell to their use. It was this system that led to John Hilliard's correction of the blurred tapping of Eleanor Powell's very rapid tap dancing by signal aligning the high and low frequency horns. They found that a 3 inch misalignment was small enough to not smear the tapping. (Late in the 1980s, I demonstrated that from 0 to 3 inch misalignment resulted in a shift in the polar response.) Hilliard had previously found that there was on the order of 1500° in phase shift in the early studio amplification systems. He corrected the problem and published his results in the 1930s.

After WWII, Hilliard and Blackburn, who both were at MIT doing radar work during the war, went their separate ways, with Hilliard joining Altec Lansing. Hilliard received an honorary Ph.D. with a degree from the Hollywood University run by Howard Termaine, the author of the original Audio Encyclopedia, the forerunner to this present volume, The Handbook for Sound Engineers.

Robert Lee Stephens left MGM in 1938 to found his own company. In the early 1950s I witnessed demonstrations of the Altec 604, the Stephens TruSonic co-axial and the Jensen Triaxial, side by side in my hi-fi shop, The Golden Ear. The Tru-Sonics was exceptionally clean and efficient. Stephens also made special 15 inch low-frequency drivers for the early Klipschorns. Hilliard, Stephens, Lansing and Shearer defined the theater loudspeaker for their era with much of the design of the Shearer multicells manufactured by Stephens.

When James Lansing (aka James Martinella) first came west he adopted the Hollywood technique of a name change. His brother, who worked for Altec through his entire career, shortened his name to Bill Martin, a truly skilled machinist who could tool anything. In 1941, Altec bought Lansing Manufacturing Company and changed the Altec name to Altec Lansing Corp. James Lansing was enjoined by Altec to the use of JBL rather than Lansing for product names. He committed suicide in 1949, and JBL would have vanished except Edmond May, considered the most valuable engineer ever at JBL, stepped into the design breach with a complete series of high quality products.

In 1947, Altec purchased Peerless Electrical Products Co. This brought in not only the first designers of 20–20,000 Hz output transformer, Ercel Harrison and his talented right-hand man, Bob Wolpert, but also the ability to manufacture what they designed. Ercel Harrison's Peerless transformers are still without peer even today.

In 1949, Altec acquired the Western Electric Sound Products Division and began producing the W.E. product lines of microphones and loudspeakers. It was said that all the mechanical product tooling, such as turntables and camera items were dumped in the channel between Los Angeles and Catalina.

Jim Noble

Jim Noble, H.S. Morris, Ercel Harrison, John Hillard, Jim Lansing, Bob Stevens and Alex Badmieff (my co-author for How to Build Speaker Enclosures) were among the giants who populated Altec and provided a glimpse into the late 1920s, the fabulous 1930s, and the final integration of W.E. Broadcasting and Recording technologies into Altec in the 1950s.

Paul Klipsch in 1959 introduced me to Art Crawford, the owner of a Hollywood FM station, who developed the original duplex speaker. The Hollywood scene has always had many clever original designers whose ideas were for "one only" after which their ideas migrated to manufacturers on the West coast.

Running parallel through the 20s and 30s with the dramatic developments by Western Electric, Bell Labs, and RCA were the entrepreneurial start-ups by men like Sidney N. Shure of Shure Brothers, Lou Burroughs and Al Kahn of what became Electro-Voice, and E. Norman Rauland who from his early Chicago radio station WENR went on to become an innovator in cathode ray tubes for radar and early television.

When I first encountered these in men in the 50s, they sold their products largely through parts pistributors. Starting the 1960s they sold to sound contractors. Stromberg-Carlson, DuKane, RCA, and Altec were all active in the rapidly expanding professional sound contractor market.

A nearly totally overlooked engineer in Altec Lansing history is Paul Veneklasen, famous in his own right for the Western Electro Acoustic Laboratory, WEAL. During WWII, Paul Veneklasen researched and designed, through extensive outdoor tests with elaborate towers, what became the Altec Voice of the Theater in postwar America. Veneklasen left Altec when this and other important work (the famed "wand" condenser microphone) were presented as Hilliard's work in Hilliard's role as a figurehead. Similar tactics were used at RCA with Harry Olson as the presenter of new technology. Peter Goldmark of the CBS Laboratories was given credit for the 33$^{1}/_{3}$ long playing record. Al Grundy was the engineer in charge of developing it, but was swept aside inasmuch as CBS used Goldmark as an icon for their introductions. Such practices were not uncommon when large companies attempted to put an "aura" around personnel who introduced their new products, to the chagrin and disgust of the actual engineers who had done the work.

"This is the west, sir, and when a legend and the facts conflict, go print the legend."
From Who Shot Liberty Valance

Audio Publications

Prior to WWII, the IRE, Institute of Radio Engineers, and the AIEE, American Institute of Electrical Engineers, were the premier sources of technology applicable to audio. The Acoustical Society of America filled the role in matters of acoustics. I am one year older than the JASA, which was first published in 1929. In 1963, the IRE and AIEE merged to become the IEEE, the Institute of Electrical and Electronic Engineers.

In 1947, C.G. McProud published Audio Engineering that featured construction articles relevant to Audio. Charles Fowler and Milton Sleeper started High Fidelity in 1954. Sleeper later published Hi Fi Music at Home. These magazines were important harbingers of the explosive growth of component sound equipment in the 1950s.

The Audio Engineering Society, AES, began publication of their journal in January 1953. The first issue contained an article written by Arthur C. Davis entitled, "Grounding, Shielding and Isolation."

Readers need to make a clear distinction in their minds between magazines designed as advertising media for "fashion design" sound products and magazines that have the necessary market information requiring the least reader screening of foolish claims. The right journals are splendid values and careful perusal of them can bring the disciplined student to the front of the envelope rapidly.

The "High" Fidelity Equipment Designers

Lincoln Walsh

By the beginning of WWII, Lincoln Walsh had designed what is still today considered the lowest distortion power amplifier using all triode 2A3s. Solid state devices, even today, have yet to match the perfection of amplifiers such as Lincoln Walsh's Brook with its all triode 2A3s or Marantz's EL34 all triode amplifier. The Walsh amplifiers, with the linearity and harmonic structure achieved by these seminal tube amplifiers, are still being constructed by devotees of fidelity who also know how to design reasonable efficiency loudspeakers. One engineer that I have a high regard for tells the story,

> *It wasn't that long ago I was sitting with the editor of a national audio magazine as his $15,000 transistor amplifier expired in a puff of smoke and took his $22,000 speakers along for the ride. I actually saw the tiny flash of light as the woofer voice coil vaporized from 30 A of dc offset—true story folks.*

In the 1950s, a group of Purdue University engineers and I compared the Brook 10 W amplifier to the then very exciting and unconventional 50 W McIntosh. The majority preferred the 10 W unit. Ralph Townsley, chief engineer at WBAA, loaned us his peak reading meter. This was an electronic marvel that weighed about 30 lbs but could read the true full peak side-by-side with the VU reading on two beautiful VI instruments. We found that the ticks on a vinyl record caused clipping on both amplifiers but the Brook handled these transients with far more grace than the McIntosh.

We later acquired a 200 W tube-type McIntosh and found that it had sufficient headroom to avoid clipping over the Klipschorns, Altec 820s, etc.

When Dr. R.A. Greiner of the University of Wisconsin published his measurements of just such effects, our little group were appreciative admirers of his extremely detailed measurements. Dr. Greiner could always be counted on for accurate, timely, and when necessary, myth-busting corrections. He was an impeccable source of truth. The home entertainment section of audio blithely ignored his devastating examination of their magical cables and went on to fortunes made on fables.

Music reproduction went through a phase of, to this writer, backward development, with the advent of extremely low efficiency book shelf loudspeaker packages with efficiencies of 20–40 dB below the figures which were common for the horn loudspeakers that dominated the home market after WWII. Interestingly, power amplifiers today are only 10–20 dB more powerful than a typical 1930s triode amplifier.

I had the good fortune to join Altec just as the fidelity home market did its best to self-destruct via totally unreliable transistor amplifiers trying to drive "sinkholes" for power loudspeakers in a marketing environment of spiffs, department store products, and the introduction of source material not attractive to trained music listeners.

I say "good fortune" as the professional sound was, in the years of the consumer hiatus, to expand and develop in remarkable ways. Here high efficiency was coupled to high power, dynamic growth in directional control of loudspeaker signals, and the growing awareness of the acoustic environment interface.

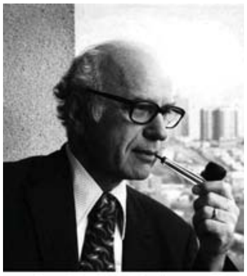
Sound System Equalization

Harry Olson

Harry Olson and John Volkmann at RCA made many advances with dynamical analogies, equalized loudspeakers, and an array of microphone designs.

Dr. Wayne Rudmose was the earliest researcher to perform meaningful sound system equalization. Dr. Rudmose published a truly remarkable paper in Noise Control (a supplementary journal of the Acoustical Society of America) in July 1958. At the AES session in the fall of 1967, I gave the first paper on the ⅓ octave contiguous equalizer. Wayne Rudmose was the chairman of the session.

In 1969, a thorough discussion of acoustic feedback that possessed absolute relevance to real-life equalization appeared in the Australian Proceedings of the IREE. "A Feedback-Mode Analyzer/Suppressor Unit for Auditorium Sound System Stabilization" by J.E. Benson and D.F. Craig, illustrating the step-function behavior of the onset and decay of regeneration in sound systems.


Dr. Wayne Rudmose

These four sources constitute the genesis of modern system equalization. Fixed equalization was employed by many early experimenters including Kellogg and Rice in the early 1920s, Volkmann of RCA in the 1930s, and Dr. Charles Boner in the 1960s.

Dr. Boner is shown here in the midst of installing filters hard-wired one at a time "until the customer ran out of money"— was a famous quote. His demonstrations of major improvements in sound systems installed in difficult environments encouraged many to further investigate sound system design and installation practices, followed by custom ⅓ octave equalization. His view of himself was "that the sound system was the heart patient and he was the Dr. DeBakey of sound."

Dr. Charles Boner

The equalization system developed at Altec in 1967 by Art Davis (of Langevin fame), Jim Noble, chief electronics engineer, and myself was named Acousta-Voicing. This program, coupled precision measurement equipment and specially trained sound contractors, resulted in larger more powerful sound systems once acoustic feedback was tamed via band rejection filters spaced at ⅓ octave centers.

Art Davis in his lab at Altec

Equalization dramatically affected quality in recording studios and motion picture studios. I introduced variable system equalization in special sessions at the screening facilities in August 1969 to the sound heads of MGM—Fred Wilson, Disney — Herb Taylor, and Al Green—Warner Bros/7 Arts.

MGM
Aug • 69
DISNEY
WB/7Arts
FRED WILSON
HERB TAYLOR
AL GREEN

Sound system equalization, room treatment such as Manfred Schroeder's Residue Diffusers designed and manufactured by Peter D'Antonio, and the signal alignment of massive arrays led to previously unheard of live sound levels in large venues.

Acoustics

As Kelvin was to electrical theory so was John William Strutt, Third Baron Rayleigh, to acoustics. He was known to later generations as Lord Rayleigh (1842–1919). I was employed by Paul W. Klipsch, a designer and manufacturer of high quality loudspeaker systems in the late 1950s. He told me to obtain and read Lord Rayleigh's *The Theory of Sound.* I did so to my immense long term benefit. This remarkable three-volume tome remains the ultimate example of what a gentleman researcher can achieve in a home laboratory. Lord Rayleigh wrote,

> *The knowledge of external things which we derive from the indications of our senses is for the most part the result of inference.*

The illusionary nature of reproduced sound, the paper cone moving back and forth being inferred to be a musical instrument, a voice, or other auditory stimuli, was vividly reinforced by the famous Chapter 1.

Wallace Clement Sabine

In terms of room acoustics, Wallace Clement Sabine was the founder of the science of architectural acoustics. He was the acoustician for Boston Symphony Hall, which is considered to be one of the three finest concert halls in the world. He was the mountain surrounded by men like Hermann, L.F. von Helmholtz, Lord Rayleigh, and others—early insights into how we listen and perceive. As one both researches and recalls from experience the movers and shakers of the audio-acoustic industry, the necessity to publish ideas is paramount.

Boston Symphony Hall

Modern communication theory has revealed to us a little of the complexity of the human listener. The human brain has from 10^{15} to 10^{17} bits of storage and we are told an operating rate of 100,000 Teraflops per second. No wonder some "sensitives" found difficulties in early digital recordings and even today attendance at a live unamplified concert quickly dispels the notion that reproduced sound has successfully modeled live sound.

We have arrived in the 21st century with not only fraudulent claims for products (an ancient art) but deliberately fraudulent technical society papers hoping to deceive the reader. I once witnessed a faulty technical article in a popular audio magazine that caused Mel Sprinkle (authority on the gain and loss of audio circuits) to write a Letter to the Editor. The Editor wrote saying Mel must be the one in error as a majority of the Letters to the Editor sided with the original author—a case of engineering democracy. We pray that no river bridges will be designed by this democratic method.

Frederick Vinton Hunt of Harvard was one of the intellectual offspring of men like Wallace Clement Sabine. As Leo Beranek wrote,

> *At Harvard, Hunt worked amid a spectacular array of physicists and engineers. There was George Washington Pierce, inventor of the crystal oscillator and of magnetostriction transducers for underwater sound; Edwin H. Hall of the Hall effect; Percy Bridgeman, Nobel Lareate, whose wife had been secretary to Wallace Sabine; A.E. Kennelly of the Kennelly-Heaviside layer; W.F. Osgood, the mathematician; O.D. Kellog of potential theory; and F.A. Saunders, who was the technical heir at Harvard to Sabine.*
>
> *Hunt's success in 1938 of producing a wide range 5 gram phonograph pickup that replaced the 5 oz units then in use led to Hunt and Beranek building large exponentially folded horns, a very high power amplifier and the introduction of much higher fidelity than had previously been available.*

Dr. Hunt attended the technical session at the Los Angeles AES meeting in 1970 when I demonstrated the computation of acoustic gain for the sound system at hand, followed by Acousta-Voicing equalization in real time on the first H.P. Real Time Analyzer, all in 20 minutes. Dr. Hunt's remark to the audience following the demonstration insured the immediate acceptance of what we had achieved without any questions from the previous doubters. Dr. Hunt took genuine interest in the technology and was generous in his praise of our application of it. He said, "I don't fully understand how you have done it, but it certainly works."

Professional-Level Audio Equipment Scaled to Home Use

World War II had two major consequences in my life (I just missed it by one year). The first was going to college with the returning G.I.s and discovering the difference in maturity between a gung-ho kid and a real veteran only one or two years older. The chasm was unbridgeable and left a lifelong respect for anyone who has served their country in the armed services.

As a young ham operator, I had obtained a very small oscilloscope, McMillan, for use as a modulation monitor. I had seen the General Radio type 525A at Purdue University, without realizing until many years later, the genius it embodied by Professor Bedell of Cornell, inventor of the linear sweep circuit, and H.H. Scott while working on it as a student at MIT with a job at General Radio as well.

The second was the pent-up explosion of talent in the audio industry especially that part misnamed *hi-fidelity.* Precision high quality it was, fidelity we have yet to achieve.

Directly after WWII a demand arose for professional level sound equipment scaled to "in the home use." Innovators such as Paul Klipsch, Lincoln Walsh, Frank McIntosh, Herman Hosmer Scott, Rudy Bozak, Avery Fisher, Saul Marantz, Alex Badmieff, Bob Stevens, and James B. Lansing met the needs of those desiring quality sound capable of reproducing the FM broadcasts and the fuller range that the advent of $33^1/_3$ vinyl records brought about.

During the early 50s, Lafayette and West Lafayette were two small towns across from each other on the banks of the Wabash River. Our clientele, Indiana's first hi-fi shop, the Golden Ear, was drawn from Purdue University and men like those named above could draw audiences equipped to appreciate their uniqueness. At that period Purdue had one of the finest minds in audio in charge of its broadcast station WBAA, Indiana's first broadcasting station and consequently a "clear channel" that Ralph Townsley utilized to modulate 20–20,000 Hz low distortion AM signals. Those of us who had Sargent Rayment TRF tuners had AM signals undistinguishable from FM, except during electrical storms. Any graduating Electrical engineer who could pass Townsley's basic audio networks test, for a job at WBAA, was indeed an engineer who could think for himself or herself about audio signals.

Atwater Kent

Great audio over AM radio in the late 1920s and early 1930s ran from the really well-engineered Atwater Kent tuned radio frequency receiver (still the best way to receive AM signals via such classics as the Sargent Rayment TRF tuner) to the absolutely remarkable, for its time, E.H. Scott's Quaranta (not to be confused with the equally famous H.H. Scott of postwar years).

This was a 48 tube superheterodyne receiver built on six chrome chassis weighing 620 lbs with five loudspeakers (two woofers, midrange, and high frequency units) biamped with 50 W for the low frequencies and 40 W for the high frequencies. My first view of one of these in the late 1930s revealed that wealth could provide a cultural life.

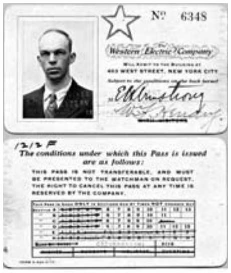
. Scott Quaranta receiver

Edwin Armstrong (1890–1954) The Invention of Radio and Fidelity

Edwin Armstrong

The technical history of radio is best realized by the inventor/engineer Edwin Howard Armstrong. Other prominent figures were political and other engineers were dwarfed by comparison to Armstrong.

In the summer of 1912, Armstrong, using the new triode vacuum tube, devised a new regenerative circuit in which part of the signal at the plate was fed back to the grid to strengthen incoming signals. In spite of his youth, Armstrong had his own pass to the famous West Street Bell Labs because of his regenerative circuit work. The regenerative circuit allowed great amplification of the received signal and also was an oscillator, if desired, making continuous wave transmission possible. This single circuit became not only the first radio amplifier, but also the first continuous wave transmitter that is still the heart of all radio operations.



Nº 6348

Western Electric Company

WILL ADMIT TO THE BUILDING AT
463 WEST STREET, NEW YORK CITY

EHArmstrong

12/12 F

The conditions under which this Pass is issued are as follows:

THIS PASS IS NOT TRANSFERABLE, AND MUST BE PRESENTED TO THE WATCHMAN ON REQUEST. THE RIGHT TO CANCEL THIS PASS AT ANY TIME IS RESERVED BY THE COMPANY.

Edwin Armstrong's pass to West Street Labs

In 1912–1913 Armstrong received his engineering degree from Columbia University, filed for a patent, and then returned to the university as assistant to professor and inventor Michael Pupin. Dr. Pupin was a mentor to Armstrong and a great teacher to generations at Columbia University.

World War I intervened and Armstrong was commissioned as an officer in the U.S. Army Signal Corps and sent to Paris. While there and in the pursuit of weak enemy wireless signals, he designed a complex eight tube receiver called the *superheterodyne circuit*, the circuit still used in 98% of all radio and television receivers.

In 1933 Armstrong invented and demonstrated wide-band frequency modulation that in field tests gave clear reception through the most violent storms and the greatest fidelity yet witnessed. The carrier was constant power while the frequency was modulated over the bandpass chosen.

Master Oscillator / Balanced Modulator

Edward Armstrong's breadboard system

He had built the entire FM transmitter and receiver on breadboard circuits of Columbia University. After the fact of physical construction, he did the mathematics.

Armstrong, in developing FM, got beyond the equations of the period which in turn laid the foundations for information theory, which quantifies how bandwidth can be exchanged for noise immunity.

In 1922, John R. Carson of AT&T had written an IRE paper that discussed modulation mathematically. He showed that FM could not reduce the station bandwidth to less than twice the frequency range of the audio signal, "Since FM could not be used to narrow the transmitted band, it was not useful."

Edwin Armstrong ignored narrowband FM and moved his experiments to 41 MHz and used a 200 kHz channel for wideband, noiseless reproduction. FM broadcasting allowed the transmitter to operate at full power all the time and used a limiter to strip off all amplitude noise in the receiver. A detector was designed to convert frequency variations into amplitude variations.

Paul Klipsch was a personal friend of Edwin Armstrong: Mr. Klipsch had supplied Klipschorns for the early FM demonstration just after WWII. This was when Armstrong, through Sarnoff's political manipulation, had been forced to move FM from 44–50 MHz to 88–108 MHz, requiring a complete redesign of all equipment. It was a stark lesson on how the courts, the media, and really big money can destroy genuine genius. Armstrong had literally created radio: the transmitters, the receivers for AM-FM-microwave in their most efficient forms. David Sarnoff made billions out of Armstrong's inventions, as well as an economic-political empire via the AM radio networks. No court or any politician should ever be allowed to make a technical judgment. Those judgments should be left to the technical societies as the "least worst" choice.

The history of audio is not the forum for discussing the violent political consequences—Sarnoff of RCA totally controlled the powerful AM networks of the time. In 1954 attorneys for RCA and AT&T led to Armstrong's death by suicide. The current AM programming quality put on FM leaves quality FM radio a rare luxury in some limited areas.

1950s music system

The few, myself included, who heard the live broadcasts of the Boston Symphony Orchestra over the FM transmitter given them by Armstrong and received on the unparalleled, even today, precedent FM receivers know what remarkable transparency can be achieved between art and technology.

Acoustic Measurements—Richard C. Heyser (1931–1987)

Richard Heyser

Plato said, "God ever geometrizes." Richard Heyser, the geometer, should feel at ease with God. To those whose minds respond to the visual, Heyser's measurements shed a bright light on difficult mathematical concepts. The Heyser Spiral displays the concepts of the complex plane in a single visual flash. Heyser was a scientist in the purest sense of the word, employed by NASA, and audio was his hobby. I am quite sure that the great scientists of the past were waiting at the door for him when he past through. His transform has yet to be fully understood. As with Maxwell, we may have to wait a hundred years.

When I first met Richard C. Heyser in the mid-1960s, Richard worked for Jet Propulsion Labs as a senior scientist. He invited me to go to his basement at his home to see his personal laboratory. The first thing he showed me on his Time Delay Spectrometry equipment was the Nyquist plot of a crossover network he was examining. I gave the display a quick look and said,

"That looks like a Nyquist plot!"

He replied, "It is."

"But," I said, "No one makes a Nyquist analyzer."

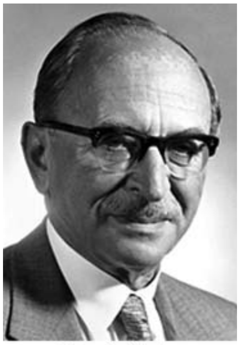
"That's right," he replied.

At this point I entered the modern age of audio analysis. Watching Dick tune in the signal delay between his microphone and the loudspeaker he was testing until the correct bandpass filter Nyquist display appeared on the screen was a revelation. Seeing the epicycles caused by resonances in the loudspeaker and the passage of non-minimum phase responses back through all quadrants opened a million questions.

Dick then showed me the Bode plots of both frequency and phase for the same loudspeaker but I was to remain a fan of seeing everything at once via the Nyquist plot.

To put all this in perspective (I worked at Altec at the time) I knew of no manufacturer in audio capable of making any of these measurements. We all had Bruel and Kjaer or General Radio frequency analyzers and good Tektronics oscilloscopes, but zero true acoustic phase measurement capabilities. I do not mean to imply that the technology didn't exist because Wente calculated the phase response of 555A in the 1920s, but rather that commercial instruments available in audio did not exist until Richard Heyser demonstrated the usefulness of the measurements and Gerald Stanley of Crown International actually built a commercially available device. Heyser's remarkable work became the Time, Envelope, Frequency (TEF) system, first in the hands of Crown International, and later as a Gold Line instrument.

The early giants of audio computed theoretical phase responses for minimum phase devices. A few pure scientists actually measured phase—Weiner, Ewask, Marivardi and Stroh, but their results had failed to go beyond their laboratories.


Dennis Gabor

From 1966 until today, 42 years later, such analysis can now be embodied in software in fast, large memory computers. Dennis Gabor's (1900–1979) analytic signal theory appeared in Heyser's work as amplitude response, phase response, and Envelope Time Curves (ETC). One glance at the Heyser Spiral for impedance reveals Gabor's analytic signal and the complex numbers as real, imaginary, and Nyquist plot. The correlation of what seems first to be separate components into one component is a revelation to the first time viewer of this display. The unwinding of the Nyquist plot along the frequency axis provides a defining perspective.

Heyser's work led to loudspeakers with vastly improved spatial response, something totally unrecognized in the amplitude-only days. Arrays became predictable and coherent. Signal alignment entered the thought of designers. The ETC technology resulted in the chance to meaningfully study loudspeaker–room interactions.

Because the most widely taught mathematical tools proceed from impulse responses, Heyser's transform is perceived "through a glass darkly." It is left in the hands of practitioners to further the research into the transient behavior of loudspeakers. The decades-long lag of academia will eventually apply the lessons of the Heyser transform to transducer signal delay and signal delay interaction.

I have always held Harry Olson of RCA in high regard because, as editor of the Audio Engineering Society Journal in 1969, he found Richard C. Heyser's original paper in the waste basket—it had been rejected by means of the idiot system of non-peer review used by the AES Journal.

Calculators and Computers

In the late 1960s, I was invited to Hewlett Packard to view a new calculator they were planning to market. I was working at this time with Arthur C. Davis (not a relative) at Altec, and Art was a friend of William Hewlett. Art had purchased one of the very first RC oscillators made in the fabled HP garage. He had used them for the audio gear that he had designed for the movie—Fantasia.

Don Davis and Tom Osborne

The 9100 calculator–computer was the first brainchild that Tom Osborne took to HP, after having been turned down by SCM, IBM, Friden and Monroe. (I purchased one; it cost me $5100. I used it to program the first acoustic design programs.) In 1966, a friend introduced Osborne to Barney Oliver at HP. After reviewing the design he asked Osborne to come back the next day to meet Dave and Bill, to which Osborne said, "Who?" After one meeting with "Dave & Bill," Osborne knew he had found a home for his 9100. Soon Bill Hewlett turned to Tom Osborne, Dave Cochran, and Tom Whitney, who worked under the direction of Barney Oliver, and said, "I want one in a tenth the volume (the 9100 was IBM typewriter size), ten times as fast, and at a tenth of the price." Later he added that he "wanted it to be a shirt pocket machine."

The first HP 35 cost $395, was 3.2 × 5.8 × 1.3 inches and weighed 9 oz with batteries. It also fit into Bill Hewlett's shirt pocket. (Bill Hewlett named the calculator the HP 35 because it had 35 keys.) Art Davis took me to lunch one day with Mr. Hewlett. Because I had been an ardent user of the HP 9100 units, I was selected to preview the HP 35 during its initial tests in Palo Alto.

In my mind, these calculators revolutionized audio education, especially for those without advanced university educations. The ability to quickly and accurately work with logarithm, trigonometric functions, complex numbers, etc., freed us from the tyranny of books of tables, slide rules, and carefully hoarded volumes such as Massa's acoustic design charts and Vegas's ten place log tables.

For the multitude of us who had experienced difficulty in engineering courses with misplaced decimal points and slide rule manipulation and extrapolation, the HP 35 released inherent talents we didn't realize we possessed. The x^y key allowed instant K numbers. The ten-place log tables became historical artifacts.

When I suggested to the then president of Altec that we should negotiate being the one to sell the HP 35s to the electronics industry (Altec then owned Allied Radio,) his reply stunned me, "We are not in the calculator business." I thought as he said it, "Neither is Hewlett Packard." His decision made it easy for me to consider leaving Altec.

I soon left Altec and started Synergetic Audio Concepts, teaching seminars in audio education. I gave each person attending a seminar an HP 35 to use during the 3-day seminar. I know that many of those attending immediately purchased an HP calculator, which changed their whole approach to audio system design. As Tom Osborne wrote, "The HP 35 and HP 65 changed the world we live in."

Since the political demise of the Soviet Union, "Mozarts-without-a-piano" have been freed to express their brilliance. Dr. Wolfgang Ahnert, from former East Germany, was enabled to use his mathematical skills with matching computer tools to dominate the audio-acoustic design market place.

The Meaning of Communication

The future of audio and acoustics stands on the shoulders of the giants we have discussed, and numerous ones that we have inadvertently overlooked. The discoverers of new and better ways to generate, distribute, and control sound will be measured consciously or unconsciously by their predecessor's standards. Fad and fundamentals will be judged eventually. Age councils that "the ancients are stealing our inventions." The uncovering of an idea new to you is as thrilling as it was to the first person to do so.

The history of audio and acoustics is the saga of the mathematical understanding of fundamental physical laws. Hearing and seeing are illusionary, restricted by the inadequacy of our physical senses. The science and art of audio and acoustics are essential to our understanding of history inasmuch as art is metaphysical (above the physical). Also art precedes science.

That the human brain processes music and art in a different hemisphere from speech and mathematics suggests the difference between information, that can be mathematically defined and communication that cannot. A message is the flawless transmission of a text. Drama, music, and great oratory cannot be flawlessly transmitted by known physical systems. For example, the spatial integrity of a great orchestra in a remarkable acoustic space is today even with our astounding technological strides only realizable by attending the live performance.

The complexity of the auditory senses defies efforts to record or transmit it faithfully.

The perception of good audio will often flow from the listener's past experience, i.e., wow and flutter really annoys musicians whereas harmonic distortion, clipping, etc., grate on an engineer's ear–mind system.

I have not written about today's highly hyped products as their history belongs to the survivors of the early 21st century. It can be hoped that someday physicists and top engineers will for some magic reason return to the development of holographic audio systems that approach fidelity.

Telecommunication technology, fiber optics, lasers, satellites, etc. have obtained worldwide audiences for both trash and treasure.

The devilish power that telecommunications has provided demagogues is frightening, but shared communication has revealed to a much larger audience the prosperity of certain ideas over others, and one can hope that the metaphysics behind progress will penetrate a majority of the minds out there.

That the audio industry's history has barely begun is evidenced every time one attends a live performance. We will, one day, look back on the neglect of the metaphysical element, perhaps after we have uncovered the parameters at present easily heard but unmeasurable by our present sciences. History awaits the ability to generate *the* sound field rather than *a* sound field. When a computer is finally offered to us that is capable of such generation, the question it must answer is,

"How does it feel?"

Bibliography

Lynn Olson, *The Soul of Sound,* revised and updated in 2005—a superb, accurate, and insightful picture of the greats and ingrates of consumer audio.

American Institute of Electrical Engineers, April 18, 1893. Ms. Mary Ann Hoffman of the IEEE History Center at Rutgers University is in the process of rescuing these seminar papers and getting them to restorers. She graciously provided me with a copy of the original.

Eugene Patronis, pictures of 1930s Western Electric theater loudspeaker.

William T. McQuaide, Audio Engineering Society, picture of Lincoln Walsh.

36 years of Syn-Aud-Con Newsletters for pictures, dates, and equipment.

Michael P. Frank, *Physical Limits of Computing,* University of Florida.

The ownership of some 750 audio and acoustic technical volumes now owned and preserved by Mary Gruska, who sensed their metaphysical value and purchased "the infinite with the finite."

Lawrence Lessing, *Man of High Fidelity, Edwin Howard Armstong*, Lippincott, 1956.

Tom Osborne, *Tom Osborne's Story in His Own Words*, a letter to Dave Packard explaining his development of the HP 9100 and the HP 35.

The vast resource of the Internet.

Chapter 2

Fundamentals of Audio and Acoustics

by Pat Brown

2.1 Introduction 23
2.2 The Decibel 23
2.3 Loudness and Level 26
2.4 Frequency 26
2.5 Wavelength 27
2.6 Surface Shapes 29
2.7 Superposition 30
2.8 Ohm's Law 32
2.9 Human Hearing 34
2.10 Monitoring Audio Program Material 35
2.11 Sound Propagation 36
  2.11.1 The Point Source 36
  2.11.2 The Line Source 37
2.12 Conclusion 38
Bibliography 39

This page intentionally left blank

2.1 Introduction

Many people get involved in the audio trade prior to experiencing technical training. Those serious about practicing audio dig in to the books later to learn the physical principles underlying their craft. This chapter is devoted to establishing a baseline of information that will prove invaluable to anyone working in the audio field.

Numerous tools exist for those who work on sound systems. The most important are the mathematical tools. Their application is independent of the type of system or its use, plus, they are timeless and not subject to obsolescence like audio products. Of course, one must always balance the mathematical approach with real-world experience to gain an understanding of the shortcomings and limitations of the formulas. Once the basics have been mastered, sound system work becomes largely intuitive.

Audio practitioners must have a general understanding of many subjects. The information in this chapter has been carefully selected to give the reader the big picture of what is important in sound systems. Many of the topics are covered in greater detail in other chapters of this book. In this initial treatment of each subject, the language of mathematics has been kept to a minimum, opting instead for word explanations of the theories and concepts. This provides a solid foundation for further study of any of the subjects. Considering the almost endless number of topics that could be included here, I selected the following based on my own experience as a sound practitioner and instructor. They are:

1. The Decibel and Levels.
2. Frequency and Wavelength.
3. The Principle of Superposition.
4. Ohm's Law and the Power Equation.
5. Impedance, Resistance, and Reactance.
6. Introduction to Human Hearing.
7. Monitoring Audio Program Material.
8. Sound Radiation Principles.
9. Wave Interference.

A basic understanding in these areas will provide the foundation for further study in areas that are of particular interest to the reader. Most of the ideas and principles in this chapter have existed for many years. While I haven't quoted any of the references verbatim, they get full credit for the bulk of the information presented here.

2.2 The Decibel

Perhaps the most useful tool ever created for audio practitioners is the decibel (dB). It allows changes in system parameters such as power, voltage, or distance to be related to level changes heard by a listener. In short, the decibel is a way to express "how much" in a way that is relevant to the human perception of loudness. We will not track its long evolution or specific origins here. Like most audio tools, it has been modified many times to stay current with the technological practices of the day. Excellent resources are available for that information. What follows is a short study on how to use the decibel for general audio work.

Most of us tend to consider physical variables in linear terms. For instance, twice as much of a quantity produces twice the end result. Twice as much sand produces twice as much concrete. Twice as much flour produces twice as much bread. This linear relationship does not hold true for the human sense of hearing. Using that logic, twice the amplifier power should sound twice as loud. Unfortunately, this is not true.

Perceived changes in the loudness and frequency of sound are based on the *percentage* change from some initial condition. This means that audio people are concerned with ratios. A given ratio always produces the same result. Subjective testing has shown that the power applied to a loudspeaker must be increased by about 26% to produce an audible change. Thus a ratio of 1.26:1 produces the minimum audible change, regardless of the initial power quantity. If the initial amount of power is 1 watt, then an increase to 1.26 watts (W) will produce a "just audible" increase. If the initial quantity is 100 W, then 126 W will be required to produce a just audible increase. A number scale can be linear with values like 1, 2, 3, 4, 5, etc. A number scale can be proportional with values like 1, 10, 100, 1000, etc. A scale that is calibrated proportionally is called a *logarithmic* scale. In fact, *logarithm* means "proportional numbers." For simplicity, base 10 logarithms are used for audio work. Using amplifier power as an example, changes in level are determined by finding the ratio of change in the parameter of interest (e.g. wattage) and taking the base 10 logarithm. The resultant number is the level change between the two wattages expressed in Bels. The base 10 logarithm is determined using a look-up table or scientific calculator. The log conversion accomplishes two things:

1. It puts the ratio on a proportional number scale that better correlates with human hearing.
2. It allows very large numbers to be expressed in a more compact form, Fig. 2-1.

The final step in the decibel conversion is to scale the Bel quantity by a factor of ten. This step converts Bels to decibels and completes the conversion process,

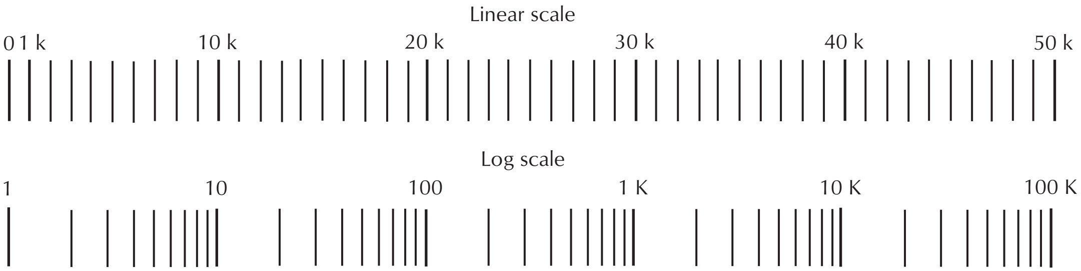

Linear scale
0 1 k
10 k
20 k
30 k
40 k
50 k
Log scale
1
10
100
1 K
10 K
100 K

Figure 2-1. A logarithmic scale has its increments marked by a fixed ratio, in this case 10 to 1, forming a more compact representation than a linear scale. Courtesy Syn-Aud-Con.

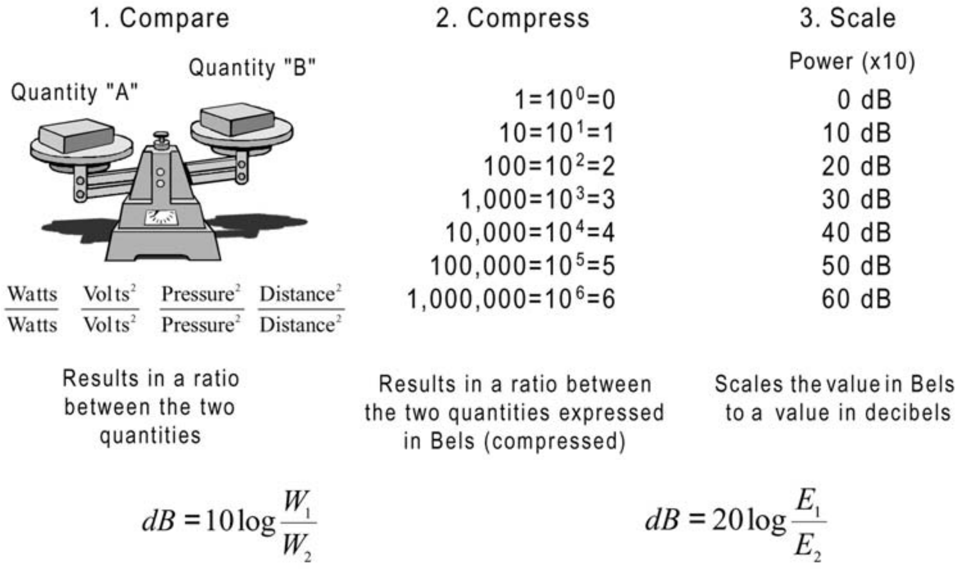

1. Compare
Quantity "A"
Quantity "B"
Watts/Watts
Volts²/Volts²
Pressure²/Pressure²
Distance²/Distance²
Results in a ratio between the two quantities
2. Compress
$1=10^0=0$
$10=10^1=1$
$100=10^2=2$
$1,000=10^3=3$
$10,000=10^4=4$
$100,000=10^5=5$
$1,000,000=10^6=6$
Results in a ratio between the two quantities expressed in Bels (compressed)
3. Scale
Power (x10)
0 dB
10 dB
20 dB
30 dB
40 dB
50 dB
60 dB
Scales the value in Bels to a value in decibels
$dB = 10\log\frac{W_1}{W_2}$
$dB = 20\log\frac{E_1}{E_2}$

Figure 2-2. The steps to performing a decibel conversion are outlined. Courtesy Syn-Aud-Con.

Fig. 2-2. The decibel scale is more resolute than the Bel scale.

The decibel is always a power-related ratio. Electrical and acoustical power changes can be converted exactly in the manner described. Quantities that are not powers must be made proportional to power—a relationship established by the power equation.

$$W = \frac{E^2}{R} \tag{2-1}$$

where,
W is power in watts,
E is voltage in volts,
R is resistance in ohms.

This requires voltage, distance, and pressure to be squared prior to taking the ratio. Some practitioners prefer to omit the squaring of the initial quantities and simply change the log multiplier from ten to twenty. This produces the same end result.

Fig. 2-3 provides a list of some dB changes along with the ratio of voltage, pressure, distance, and power required to produce the indicated dB change. It is a worthwhile endeavor to memorize the changes indicated in bold type and be able to recognize them by listening.

A decibel conversion requires two quantities that are in the same unit, i.e., watts, volts, meters, feet. The unit cancels during the initial division process, leaving the ratio between the two quantities. For this reason, the decibel is without dimension and is therefore technically not a unit in the classical sense. If two arbitrary quantities of the same unit are compared, the result is a relative level change. If a standard reference quantity is used in the denominator of the ratio, the result is an absolute level and the unit is dB relative to the original

Subjective Change	Voltage, Distance, Pressure Ratio	% of Original	PowerRatio	dB Change
	20log		**10log**	
Barely perceptible	**1.12:1**	**89**	**1.26:1**	**1 dB**
	1.26:1	79	1.58:1	2 dB
Noticeable to most	**1.41:1**	**71**	**2:1**	**3 dB**
	1.58:1	63	2.51:1	4 dB
	1.78:1	56	3.16:1	5 dB
Goal for system changes	**2:1**	**50**	**4:1**	**6 dB**
	2.24:1	45	5:1	7 dB
	2.51:1	40	6.3:1	8 dB
	2.8:1	36	8:1	9 dB
Twice as loud or soft	**3.16:1**	**32**	**10:1**	**10 dB**
	10:1	10	100:1	20 dB
	31.6:1	3	1000:1	30 dB
Limits of audibility	**100:1**	**1**	**10,000:1**	**40 dB**
	316:1	0.3	100,000:1	50 dB
	1000:1	0.1	1,000,000:1	60 dB

Figure 2-3. Some important decibel changes and the ratios of power, voltage, pressure, and distance that produce them. Courtesy Syn-Aud-Con.

unit. Relative levels are useful for live work. Absolute levels are useful for equipment specifications and calibration. Fig. 2-4 lists some references used for determining absolute levels.

The decibel was originally used in impedance-matched interfaces and always with a power reference. Power requires knowledge of the resistance that a voltage is developed across. If the resistance value is fixed, changes in applied voltage can be expressed in dB, since the power developed will be directly proportional to the applied voltage. In modern sound systems, few device interfaces are impedance matched. They are actually mismatched to optimize the voltage transfer between components. While the same impedance does not exist at each device interface, the same impedance condition may. If a minimum 1:10 ratio exists between the output impedance and input impedance, then the voltage transfer is essentially independent of the actual output or input impedance values. Such an interface is termed *constant voltage*, and the signal source is said to be operating *open circuit* or *un-terminated*. In constant voltage interfaces, open circuit conditions are assumed when using the decibel. This means that the level change at the output of the system is caused by changing the voltage somewhere in the processing chain and is dependent on the voltage change only, not the resistance that it is developed across or the power transfer. Since open-circuit conditions exist almost universally in modern analog systems, the practice of using the decibel with a voltage reference is widespread and well-accepted.

Electrical Power	
dBW	1 Watt
dBm	0.001 Watt
Acoustical Power	
dB-PWL or L_w	10^{-12} Watt
Electrical Voltage	
dBV	1 Volt
dBu	0.775 Volts
Acoustical Pressure	
dB SPL or L_p	0.00002 Pascals

Figure 2-4. Some common decibel references used by the audio industry.

One of the major utilities of the decibel is that it provides a common denominator when considering level changes that occur due to voltage changes at various points in the signal chain. By using the decibel, changes in sound level at a listener position can be determined from changes in the output voltage of any device ahead of the loudspeaker. For instance, a doubling of the microphone output voltage produces a 6 dB increase in output level from the microphone, mixer, signal processor, power amplifier, and ultimately the sound level at the listener. This relationship assumes linear operating conditions in each device. The 6 dB increase in level from the microphone could be caused by the talker speaking 6 dB louder or by simply reducing the miking distance by one-half (a 2:1 distance ratio). The level controls on audio devices are normally calibrated in relative dB. Moving a fader by 6 dB causes the output voltage of the device (and system) to increase by a factor of 2 and the output power from the device (and system) to be increased by a factor of four.

Absolute levels are useful for rating audio equipment. A power amplifier that can produce 100 watts of continuous power is rated at

$$\begin{aligned} L_{out} &= 10\log W \\ &= 10\log 100 \\ &= 20 \text{ dBW} \end{aligned} \tag{2-2}$$

This means that the amplifier can be 20 dB louder than a 1 watt amplifier. A mixer that can output 10 volts prior to clipping can be rated at

$$\begin{aligned} L_{out} &= 20\log E \\ &= 20\log 10 \\ &= 20 \text{ dBV} \end{aligned} \qquad (2\text{-}3)$$

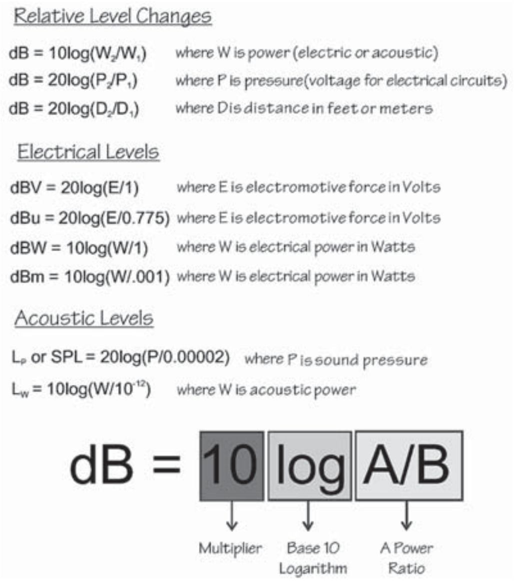
If the same mixer outputs 1 volt rms at meter zero, then the mixer has 20 dB of peak room above meter zero.

If a loudspeaker can produce a sound level at 1 meter of 90 dB ref. 20 μPa (micro-Pascals), then at 10 meters its level will be

$$\begin{aligned} L_p &= 90 + 20\log\frac{1}{10} \\ &= 90 + (-20) \\ &= 70 \text{ dB} \end{aligned} \qquad (2\text{-}4)$$

In short, the decibel says, "The level difference caused by changing a quantity will depend upon the initial value of the quantity and the *percentage* that it is changed."

The applications of the decibel are endless, and the utility of the decibel is self-evident. It forms a bridge between the amount of change of a physical parameter and the loudness change that is perceived by the human listener. The decibel is the language of audio, Fig. 2-5.



Relative Level Changes

$dB = 10\log(W_2/W_1)$ where W is power (electric or acoustic)

$dB = 20\log(P_2/P_1)$ where P is pressure(voltage for electrical circuits)

$dB = 20\log(D_2/D_1)$ where D is distance in feet or meters

Electrical Levels

$dBV = 20\log(E/1)$ where E is electromotive force in Volts

$dBu = 20\log(E/0.775)$ where E is electromotive force in Volts

$dBW = 10\log(W/1)$ where W is electrical power in Watts

$dBm = 10\log(W/.001)$ where W is electrical power in Watts

Acoustic Levels

L_p or $SPL = 20\log(P/0.00002)$ where P is sound pressure

$L_w = 10\log(W/10^{-12})$ where W is acoustic power

$dB = 10 \log A/B$

Figure 2-5. Summary of decibel formulas for general audio work. Courtesy Syn-Aud-Con.

2.3 Loudness and Level

The perceived loudness of a sound event is related to its acoustical level, which is in turn related to the electrical level driving the loudspeaker. Levels are electrical or acoustical pressures or powers expressed in decibels. In its linear range of operation, the human hearing system will perceive an increase in level as an increase in loudness. Since the eardrum is a pressure sensitive mechanism, there exists a threshold below which the signal is distinguishable from the noise floor. This threshold is about 20 μPa of pressure deviation from ambient at midrange frequencies. Using this number as a reference and converting to decibels yields

$$\begin{aligned} L_p &= 20\log\frac{0.00002}{0.00002} \\ &= 0 \text{ dB (or 0 dB SPL)} \end{aligned} \qquad (2\text{-}5)$$

This is widely accepted as the threshold of hearing for humans at mid-frequencies. Acoustic pressure levels are always stated in dB ref. 0.00002 Pa. Acoustic power levels are always stated in dB ref. 1 pW (picowatt or 10^{-12} W). Since it is usually the pressure level that is of interest, we must square the Pascals term in the decibel conversion to make it proportional to power. Sound pressure levels are measured using sound level meters with appropriate ballistics and weighting to emulate human hearing. Fig. 2-6 shows some typical sound pressure levels that are of interest to audio practitioners.

2.4 Frequency

Audio practitioners are in the wave business. A wave is produced when a medium is disturbed. The medium can be air, water, steel, the earth, etc. The disturbance produces a fluctuation in the ambient condition of the medium that propagates as a wave that radiates outward from the source of the disturbance. If one second is used as a reference time span, the number of fluctuations above and below the ambient condition per second is the frequency of the event, and is expressed in cycles per second, or Hertz. Humans can hear frequencies as low as 20 Hz and as high as 20,000 Hz (20 kHz). In an audio circuit the quantity of interest is usually the electrical voltage. In an acoustical circuit it is the air pressure deviation from ambient atmospheric pressure. When the air pressure fluctuations have a frequency between 20 Hz and 20 kHz they are audible to humans.

As stated in the decibel section, humans are sensitive to proportional changes in power, voltage, pressure, and distance. This is also true for frequency. If we start at

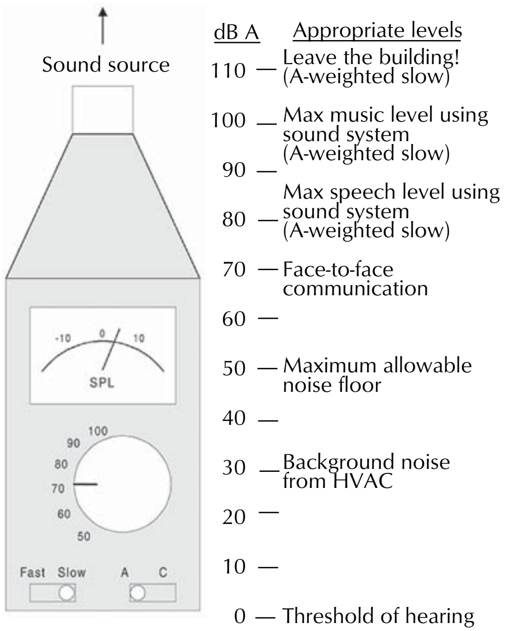

Sound source
dB A
Appropriate levels
110 — Leave the building! (A-weighted slow)
100 — Max music level using sound system (A-weighted slow)
90 —
80 — Max speech level using sound system (A-weighted slow)
70 — Face-to-face communication
60 —
50 — Maximum allowable noise floor
40 —
30 — Background noise from HVAC
20 —
10 —
0 — Threshold of hearing
-10
0
10
SPL
100
90
80
70
60
50
Fast
Slow
A
C

Figure 2-6. Sound levels of interest to system designers and operators. Courtesy Syn-Aud-Con.

the lowest audible frequency of 20 Hz and increase it by a 2:1 ratio, the result is 40 Hz, an interval of one octave. Doubling 40 Hz yields 80 Hz. This is also a one-octave span, yet it contains twice the frequencies of the previous octave. Each successive frequency doubling yields another octave increase and each higher octave will have twice the spectral content of the one below it. This makes the logarithmic scale suitable for displaying frequency. Figs. 2-7 and 2-8 show a logarithmic frequency scale and some useful divisions. The perceived midpoint of the spectrum for a human listener is about 1 kHz. Some key frequency ratios exist:

- 10:1 ratio—decade.
- 2:1 ratio—octave.

The spectral or frequency response of a system describes the frequencies that can pass through that system. It must always be stated with an appropriate tolerance, such as ±3 dB. This range of frequencies is the bandwidth of the system. All system components have a finite bandwidth. Sound systems are usually bandwidth limited for reasons of stability and loudspeaker protection. A spectrum analyzer can be used to observe the spectral response of a system or system component.

2.5 Wavelength

If the frequency f of a vibration is known, the time period T for one cycle of vibration can be found by the simple relationship

$$T = \frac{1}{f} \tag{2-6}$$

The time period T is the inverse of the frequency of vibration. The period of a waveform is the time length of one complete cycle, Fig. 2-9. Since most waves propagate or travel, if the period of the wave is known, its physical size can be determined with the following equation if the speed of propagation is known:

$$\lambda = Tc \tag{2-7}$$

$$\lambda = \frac{c}{f} \tag{2-8}$$

Waves propagate at a speed that is dependent on the nature of the wave and the medium that it is passing through. The speed of the wave determines the physical size of the wave, called its *wavelength*. The speed of light in a vacuum is approximately 300,000,000 meters per second (m/s). The speed of an electromagnetic wave in copper wire is somewhat less, usually 90% to 95% of the speed of light. The fast propagation speed of electromagnetic waves makes their wavelengths extremely long at audio frequencies, Fig. 2-10.

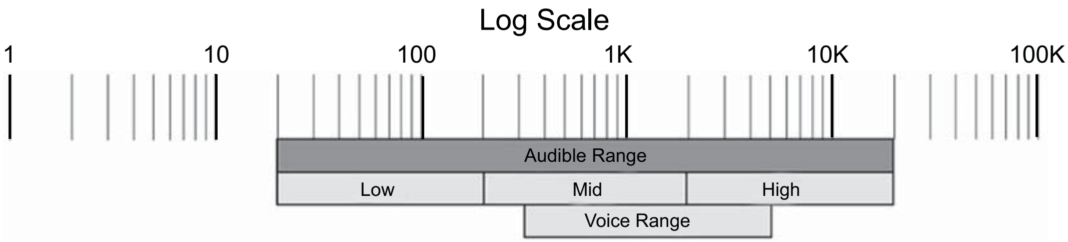

Log Scale
1
10
100
1K
10K
100K
Audible Range
Low
Mid
High
Voice Range

Figure 2-7. The audible spectrum divided into decades (a 10 to 1 frequency ratio). Courtesy Syn-Aud-Con.

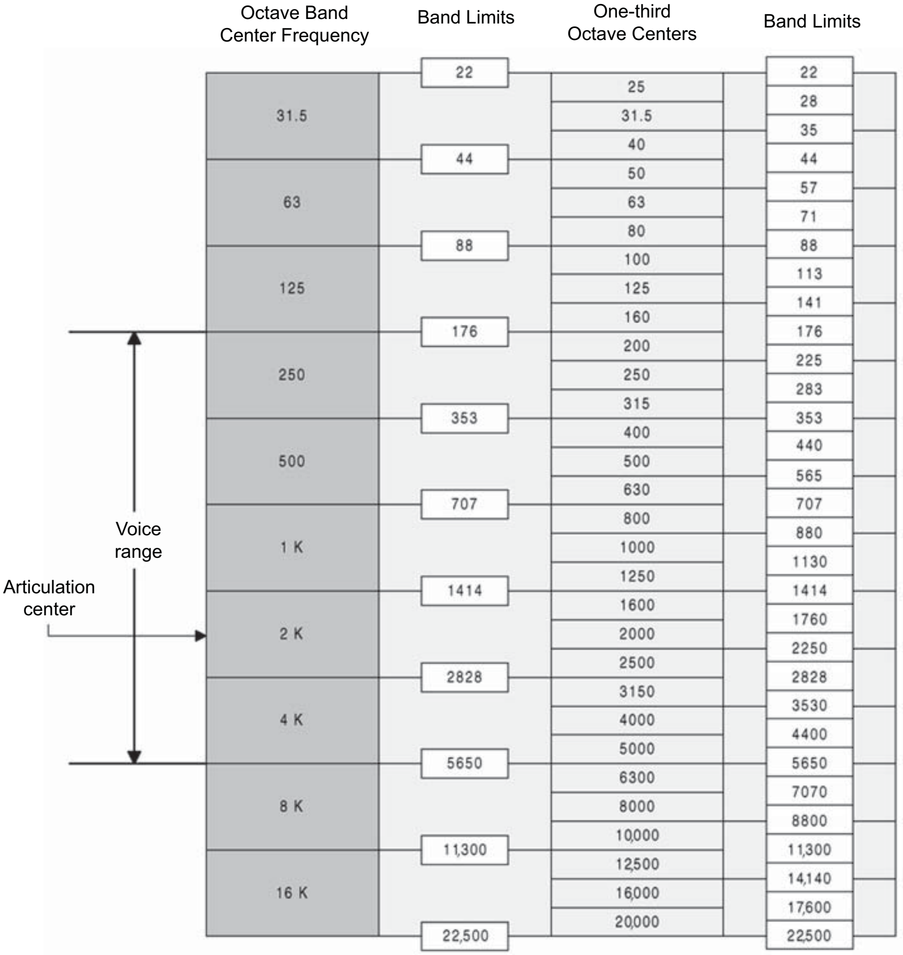

Octave Band Center Frequency
Band Limits
One-third Octave Centers
Band Limits
31.5
63
125
250
500
1 K
2 K
4 K
8 K
16 K
22
44
88
176
353
707
1414
2828
5650
11,300
22,500
25
31.5
40
50
63
80
100
125
160
200
250
315
400
500
630
800
1000
1250
1600
2000
2500
3150
4000
5000
6300
8000
10,000
12,500
16,000
20,000
22
28
35
44
57
71
88
113
141
176
225
283
353
440
565
707
880
1130
1414
1760
2250
2828
3530
4400
5650
7070
8800
11,300
14,140
17,600
22,500
Voice range
Articulation center

Figure 2-8. The audible spectrum divided into octaves (a 2 to 1 ratio) and one-third octaves. Courtesy Syn-Aud-Con.

At the higher radio frequencies (VHF and UHF), the wavelengths become very short—1 meter or less. Antennas to receive such waves must be of comparable physical size, usually one-quarter to one-half wavelength. When waves become too short for practical antennae, concave dishes can be used to collect the waves. It should be pointed out that the highest frequency that humans can hear (about 20 kHz) is a very low frequency when considering the entire electromagnetic spectrum.

An acoustic wave is one that is propagating by means of vibrating a medium such as steel, water, or air. The propagation speeds through these media are relatively slow, resulting in waves that are long in length compared to an electromagnetic wave of the same frequency. The wavelengths of audio frequencies in air range from about 17 m (20 Hz) to 17 mm (20 kHz). The wavelength of 1 kHz in air is about 0.334 m (about 1.13 ft).

When physically short acoustic waves are radiated into large rooms, there can be adverse effects from reflections. Acoustic reflections occur when a wave encounters a change in acoustic impedance, usually from a rigid surface, the edge of a surface or some other obstruction. The reflection angle equals the incidence angle in the ideal case. Architectural acoustics is the study of the behavior of sound waves in enclosed spaces. Acousticians specialize in creating spaces with reflected sound fields that enhance rather than detract from the listening experience.

When sound encounters a room surface, a complex interaction takes place. If the surface is much larger than the wavelength, a reflection occurs and an acoustic shadow is formed behind the boundary.

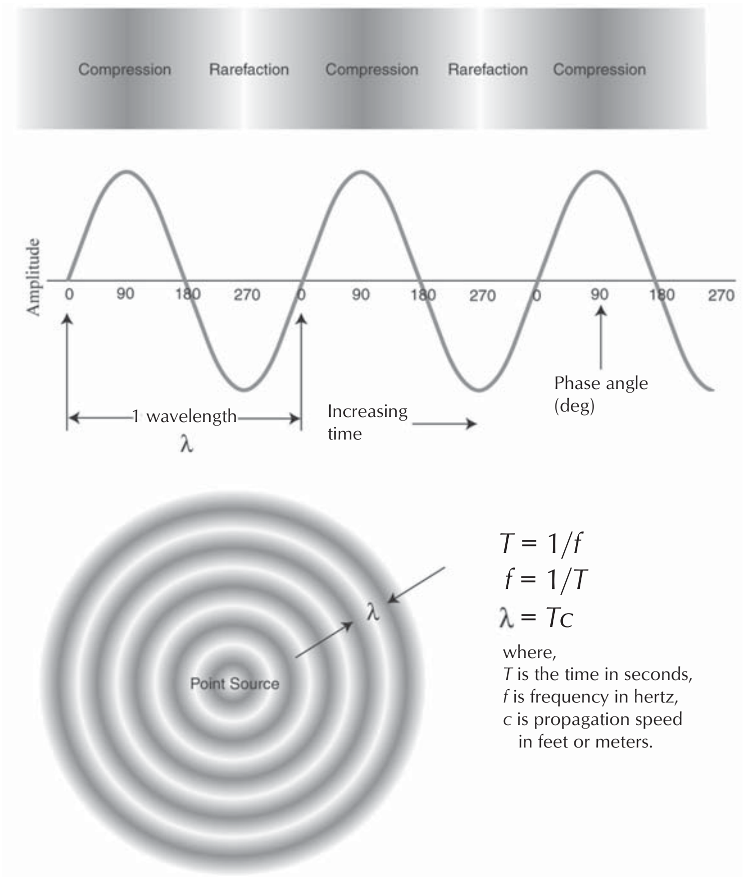

Compression
Rarefaction
Compression
Rarefaction
Compression
Amplitude
0
90
180
270
0
90
180
270
0
90
180
270
1 wavelength
λ
Increasing time
Phase angle (deg)
Point Source
λ
T = 1/f
f = 1/T
λ = Tc
where,
T is the time in seconds,
f is frequency in hertz,
c is propagation speed in feet or meters.

Figure 2-9. The wavelength of an event determines how it interacts with the medium that it is passing through. Courtesy Syn-Aud-Con.

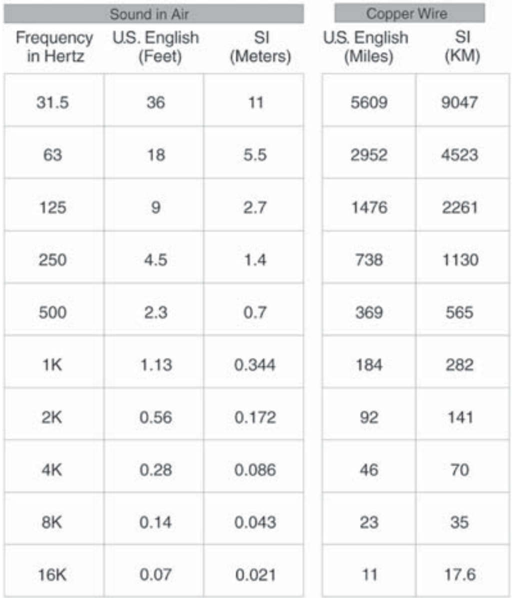
If the obstruction is smaller than the wavelength of the wave striking it, the wave diffracts around the obstruction and continues to propagate. Both effects are complex and frequency (wavelength) dependent, making them difficult to calculate, Fig. 2-11.

The reflected wave will be strong if the surface is large and has low absorption. As absorption is increased, the level of the reflection is reduced. If the surface is random, the wave can be scattered depending on the size relationship between the wave and the surface relief. Commercially available diffusors can be used to achieve uniform scattering in critical listening spaces, Fig. 2-12.

2.6 Surface Shapes

The geometry of a boundary can have a profound affect on the behavior of the sound that strikes it. From a sound reinforcement perspective, it is usually better to scatter sound than to focus it. A concave room boundary should be avoided for this reason, Fig. 2-13. Many auditoriums have concave rear walls and balcony faces that require extensive acoustical treatment for reflection control. A convex surface is more desirable, since it scatters sound waves whose wavelengths are small relative to the radius of curvature. Room corners can provide useful directivity control at low frequencies, but at high frequencies can produce problematic reflections.

Electrical reflections can occur when an electromagnetic wave encounters a change in impedance. For such waves traveling down a wire, the reflection is back towards the source of the wave. Such reflections are not usually a problem for analog waves unless there is a phase offset between the outgoing and reflected waves. Note that an audio cable would need to be very long for its length to cause a significant time offset between the



Sound in Air			Copper Wire	
Frequency in Hertz	U.S. English (Feet)	SI (Meters)	U.S. English (Miles)	SI (KM)
31.5	36	11	5609	9047
63	18	5.5	2952	4523
125	9	2.7	1476	2261
250	4.5	1.4	738	1130
500	2.3	0.7	369	565
1K	1.13	0.344	184	282
2K	0.56	0.172	92	141
4K	0.28	0.086	46	70
8K	0.14	0.043	23	35
16K	0.07	0.021	11	17.6

Figure 2-10. Acoustic wavelengths are relatively short and interact dramatically with their environment. Audio wavelengths are extremely long, and phase interaction on audio cables is not usually of concern. Courtesy Syn-Aud-Con.

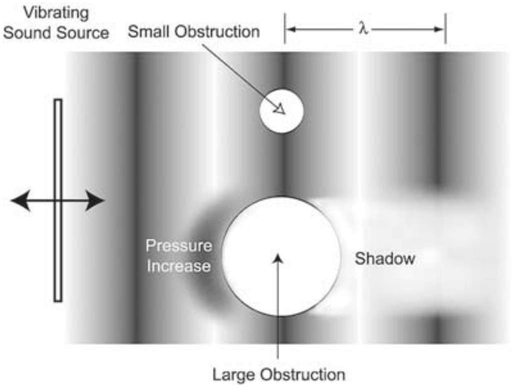

Vibrating Sound Source
Small Obstruction
λ
Pressure Increase
Shadow
Large Obstruction

Figure 2-11. Sound diffracts around objects that are small relative to the length of the sound wave. Courtesy Syn-Aud-Con.

incident and reflected wave (many thousands of meters). At radio frequencies, reflected waves pose a huge problem, and cables are normally terminated (operated into a matched impedance) to absorb the incident wave at the receiving device and reduce the level of the reflection. The same is true for digital signals due to their very high frequency content.

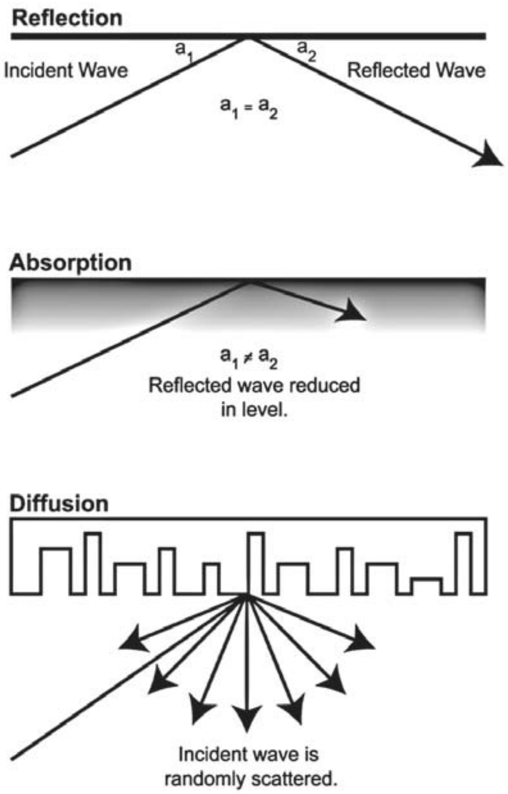
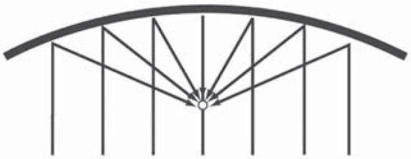

Reflection
Incident Wave
a_1
a_2
Reflected Wave
$a_1 = a_2$
Absorption
$a_1 \neq a_2$
Reflected wave reduced in level.
Diffusion
Incident wave is randomly scattered.

Figure 2-12. Sound waves will interact with a large boundary in a complex way. Courtesy Syn-Aud-Con.

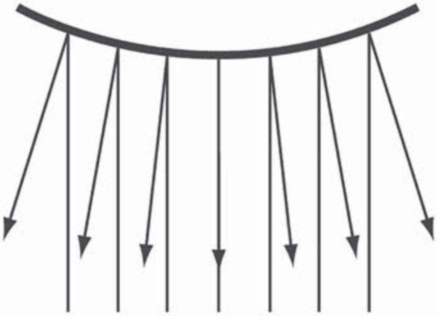
2.7 Superposition

Sine waves and cosine waves are periodic and singular in frequency. These simple waveforms are the building blocks of the complex waveforms that we listen to every day. The amplitude of a sine wave can be displayed as a function of time or as a function of phase rotation, Fig. 2-14. The sine wave will serve as an example for the following discussion about superposition. Once the size (wavelength) of a wave is known, it is useful to subdivide it into smaller increments for the purpose of tracking its progression through a cycle or comparing its progression with that of another wave. Since the sine wave describes a cyclic (circular) event, one full cycle is represented by 360°, at which point the wave repeats.

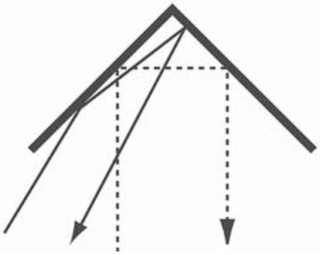
When multiple sound pressure waves pass by a point of observation, their responses sum to form a composite wave. The composite wave is the complex combination of two or more individual waves. The amplitude of the



Concave surfaces focus sound



Convex surfaces scatter sound



Corners return sound to its source

Figure 2-13. Some surfaces produce focused reflections. Courtesy Syn-Aud-Con.

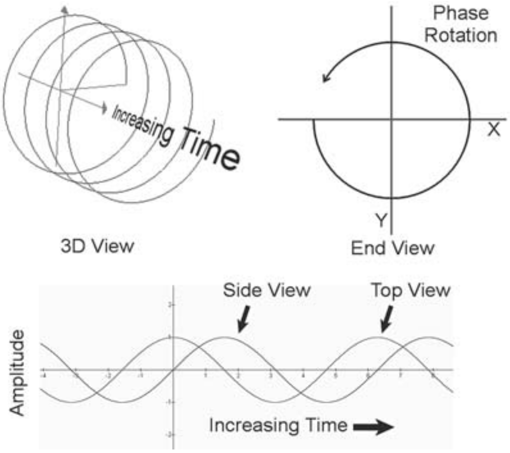


Figure 2-14. Simple harmonic motion can be represented with a sine or cosine wave. Both are viewpoints of the same event from different angles. Courtesy Syn-Aud-Con.

summation is determined by the relative phase of the individual waves. Let's consider how two waves might combine at a point of observation. This point might be a listener seat or microphone position. Two extremes exist. If there is no phase offset between two waves of the same amplitude and frequency, the result is a coherent summation that is twice the amplitude of either individual wave (+6 dB). The other extreme is a 180° phase offset between the waves. This results in the complete cancellation of the pressure response at the point of observation. An infinite number of intermediate conditions occur between these two extremes. The phase interaction of waves is not a severe problem for analog audio signals in the electromagnetic domain for sound systems, where the wavelengths at audio frequencies are typically much longer than the interconnect cables. Waves reflected from receiver to source are in phase and no cancellation occurs. This is not the case for video, radio frequency, and digital signals. The shorter wavelengths of these signals can be dramatically affected by wave superposition on interconnect cables. As such, great attention must be given to the length and terminating impedance of the interconnect cables to assure efficient signal transfer between source and receiver. The practice of impedance matching between source, cable, and load is usually employed.

In sound reinforcement systems, phase interactions are typically more problematic for acoustical waves than electromagnetic waves. Phase summations and cancellations are the source of many acoustical problems experienced in auditoriums. Acoustic wavelengths are often short relative to the room size (at least at high frequency), so the waves tend to bounce around the room before decaying to inaudibility. At a listener position, the reflected waves "superpose" to form a complex waveform that is heard by the listener. The sound radiated from multiple loudspeakers will interact in the same manner, producing severe modifications in the radiated sound pattern and frequency response. Antenna designers have historically paid more attention to these interactions than loudspeaker designers, since there are laws that govern the control of radio frequency emissions. Unlike antennas, loudspeakers are usually broadband devices that cover one decade or more of the audible spectrum. For this reason, phase interactions between multiple loudspeakers never result in the complete cancellation of sound pressure, but rather cancellation at some frequencies and coherent summation at others. The subjective result is *tonal coloration* and *image shift* of the sound source heard by the listener. The significance of this phenomenon is application-dependent. People having dinner in a restaurant would not be concerned with the effects of such interactions since they came for the food and not the music. Concert-goers or church attendees would be more concerned, because their seat might be in a dead spot, and the interactions disrupt their listening experience, possibly to the point of reducing the information conveyed via the sound system. A venue owner may make a significant investment in good quality loudspeakers, only to have their response impaired by such interactions with an adjacent loudspeaker or room surface, Fig. 2-16.

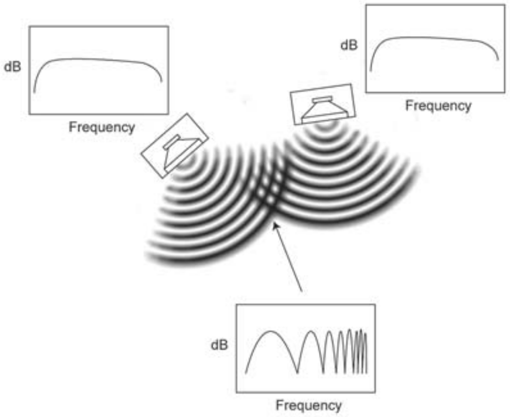


Figure 2-15. Phase interference occurs when waves from multiple sources arrive at different times. Courtesy Syn-Aud-Con.

Phase interactions are most disruptive in critical listening environments, such as recording studio control rooms or high quality home entertainment systems. Users of these types of systems often make a large investment to maintain sonic accuracy by purchasing phase coherent loudspeakers and appropriate acoustical treatments for the listening space. The tonal coloration caused by wave interference may be unacceptable for a recording studio control room but may be artistically pleasing in a home audio system.

Loudspeaker designers can use wave interaction to their advantage by choosing loudspeaker spacings that form useful radiation patterns. Almost all pattern control in the low frequency decade is achieved in this manner. Uninformed system designers create undesirable radiation patterns by accident in the way that they place and stack loudspeakers. The results are poor coverage and reduced acoustic gain.

The proper way to view the loudspeaker and room are as filters that the sound energy must pass through en route to the listener. Some aspects of these filters can be compensated with electronic filters—a process known as equalization. Other aspects cannot, and electronic *equalization* merely aggravates or masks the problem.

2.8 Ohm's Law

In acoustics, the sound that we hear is nature restoring an equilibrium condition after an atmospheric disturbance. The disturbance produces waves that cause the atmospheric pressure to oscillate above and below ambient pressure as they propagate past a point of observation. The air always settles to its ambient state upon cessation of the disturbance.

In an electrical circuit, a potential difference in electrical pressure between two points causes current to flow. Electrical current results from electrons flowing to a point of lower potential. The electrical potential difference is called an *electromotive force* (EMF) and the unit is the volt (V). The rate of electron flow is called current and the unit is the *ampere* (A). The ratio between voltage and current is called the *resistance* and the unit is the *ohm* (Ω). The product of voltage and current is the *apparent power*, W, that is produced by the source and consumed by the load. Power is the rate of doing work and power ratings must always include a reference to time. A power source can produce a rated voltage at a rated flow of current into a specified load for a specified period of time. The ratio of voltage to current can be manipulated to optimize a source for a specific task. For instance, current flow can be sacrificed to maximize voltage transfer. When a device is called upon to deliver appreciable current, it is said to be operating under load. The load on an automobile increases when it must maintain speed on an uphill grade, and greater power transfer between the engine and drive train is required. Care must be taken when loading audio components to prevent distortion or even damage. Ohm's Law describes the ratios that exist between voltage, current, and resistance in an electrical circuit.

$$R = \frac{E}{I} \tag{2-9}$$

$$E = IR \tag{2-10}$$

$$I = \frac{E}{R} \tag{2-11}$$

where,
E is in volts,
I is in amperes,
R is in ohms.

Direct current (dc) flows in one direction only. In ac (alternating current) the direction of current flow is alternating at the frequency of the waveform. Voltage and current are not always in sync so the phase relationship between them must be considered. Power flow is reduced when they are not in relative phase (synchronization). Voltage and current are in phase in resistive circuits. Phase *shifts* between voltage and current are produced by reactive elements in a circuit. Reactance reduces the power transferred to the load by storing energy and reflecting it back to the source. Loudspeakers and transformers are examples of sound system components that can have significant reactive characteristics. The combined opposition to current flow caused by resistance and reactance is termed the *impedance* (Z) of the circuit. The unit for impedance is also the ohm (Ω). An impedance can be purely resistive, purely reactive, or most often some combination of the two. This is referred to as a *complex* impedance. Impedance is a function of frequency, and impedance measurements must state the frequency at which the measurement was made. Sound system technicians should be able to measure impedance to verify proper component loading, such as at the amplifier/loudspeaker interface.

$$Z = \sqrt{R^2 + (X_T)^2} \qquad (2\text{-}12)$$

where,
Z is the impedance in ohms,
R is the resistance in ohms,
X_T is the total reactance in ohms.

Reactance comes is two forms. Capacitive reactance causes the voltage to lag the current in phase. Inductive reactance causes the current to lag the voltage in phase. The total reactance is the sum of the inductive and capacitive reactance. Since they are different in sign one can cancel the other, and the resultant phase angle between voltage and current will be determined by the dominant reactance.

In mechanics, a spring is a good analogy for capacitive reactance. It stores energy when it is compressed and returns it to the source. In an electrical circuit, a capacitor opposes changes in the applied voltage. Capacitors are often used as filters for passing or rejecting certain frequencies or smoothing ripples in power supply voltages. Parasitic capacitances can occur when conductors are placed in close proximity.

$$X_C = \frac{1}{2\pi f C} \qquad (2\text{-}13)$$

where,
f is frequency in hertz,
C is capacitance in farads,
X_C is the capacitive reactance in ohms.

In mechanics, a moving mass is analogous to an inductive reactance in an electrical circuit. The mass tends to keep moving when the driving force is removed. It has therefore stored some of the applied energy. In electrical circuits, an inductor opposes a change in the current flowing through it. As with capacitors, this property can be used to create useful filters in audio systems. Parasitic inductances can occur due to the ways that wires are constructed and routed.

$$X_L = 2\pi f L \qquad (2\text{-}14)$$

where,
X_L is the inductive reactance in ohms.

Inductive and capacitive reactance produce the opposite effect, so one can be used to compensate for the other. The total reactance X_T is the sum of the inductive and capacitive reactance.

$$X_T = X_L - X_C \qquad (2\text{-}15)$$

Note that the equations for capacitive and inductive reactance both include a frequency term. Impedance is therefore frequency dependent, meaning that it changes with frequency. Loudspeaker manufacturers often publish impedance plots of their loudspeakers. The impedance of interest from this plot is usually the *nominal* or *rated* impedance. Several standards exist for determining the rated impedance from the impedance plot, Fig. 2-16.

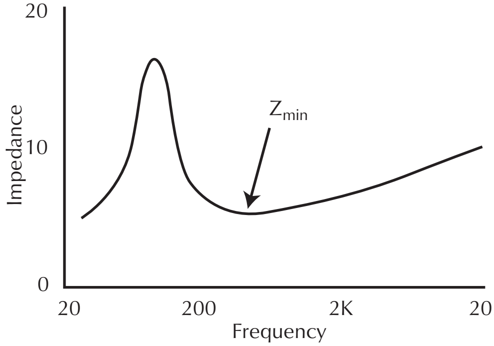


Figure 2-16. An impedance magnitude plot displays impedance as a function of the applied frequency. Courtesy Syn-Aud-Con.

An impedance phase plot often accompanies an impedance magnitude plot to show whether the loudspeaker load is resistive, capacitive, or inductive at a given frequency. A resistive load will convert the applied power into heat. A reactive load will store and reflect the applied power. Complex loads, such as loudspeakers, do both. When considering the power delivered to the loudspeaker, the impedance *Z* is used in the power equation. When considering the power dissipated by the load, the resistive portion of the impedance must be used in the power equation. The power factor describes the reduction in power transfer caused by the phase angle between voltage and current in a reactive load. Some definitions are useful.

$$\text{Apparent Power (Total Power)} = \frac{E^2}{Z} \tag{2-16}$$

$$\text{Active Power (Absorbed Power)} = \frac{E^2}{R} \tag{2-17}$$

$$\text{Reactive Power (Reflected Power)} = \frac{E^2}{Z\cos\theta} \tag{2-18}$$

where,
θ is the phase angle between the voltage and current.

Ohm's Law and the power equation in its various forms are foundation stones of the audio field. One can use these important tools for a lifetime and not exhaust their application to the electrical and acoustical aspects of the sound reinforcement system.

2.9 Human Hearing

It is beneficial for sound practitioners to have a basic understanding of the way that people hear and perceive sound. The human auditory system is an amazing device, and it is quite complex. Its job is to transduce fluctuations in the ambient atmospheric pressure into electrical signals that will be processed by the brain and perceived as sound by the listener. We will look at a few characteristics of the human auditory system that are of significance to audio practitioners.

The dynamic range of a system describes the difference between the highest level that can pass through the system and its noise floor. The threshold of human hearing is about 0.00002 Pascals (Pa) at mid frequencies. The human auditory system can withstand peaks of up to 200 Pa at these same frequencies. This makes the dynamic range of the human auditory system approximately

$$\begin{aligned} DR &= 20\log\frac{200}{0.00002} \\ &= 140\text{ dB} \end{aligned} \tag{2-19}$$

The hearing system can not take much exposure at this level before damage occurs. Speech systems are often designed for 80 dB ref. 20 μPa and music systems about 90 dB ref. 20 μPa for the mid-range part of the spectrum.

Audio practitioners give much attention to achieving a flat spectral response. The human auditory system is not flat and its response varies with level. At low levels, its sensitivity to low frequencies is much less than its sensitivity to mid-frequencies. As level increases, the difference between low- and mid-frequency sensitivity is less, producing a more uniform spectral response. The classic equal loudness contours, Fig. 2-17, describe this phenomenon and have given us the weighting curves, Fig. 2-18, used to measure sound levels.

Modern sound systems are capable of producing very high sound pressure levels over large distances. Great care must be taken to avoid damaging the hearing of the audience.

The time response of the hearing system is slow compared to the number of audible events that can occur in a given time span. As such, our hearing system *integrates* closely spaced sound arrivals (within about 35 ms) with regard to level. This is what makes sound indoors appear louder than sound outdoors. While reflected sound increases the perceived level of a sound source, it also adds colorations. This is the heart of how we perceive acoustic instruments and auditoriums. A good recording studio or concert hall produces a musically pleasing reflected sound field to a listener position. In general, secondary energy arrivals pose problems if they arrive earlier than 10 ms (severe tonal coloration) after the first arrival or later than 50 ms (potential echo), Fig. 2-19.

The integration properties of the hearing system make it less sensitive to impulsive sound events with regard to level. Peaks in audio program material are often 20 dB or more higher in level than the perceived loudness of the signal. Program material that measures 90 dBA (slow response) may contain short term events at 110 dBA or more, so care must be taken when exposing musicians and audiences to high powered sound systems.

The eardrum is a pressure sensitive diaphragm that responds to fluctuations in the ambient atmospheric pressure. Like a loudspeaker and microphone, it has an overload point at which it distorts and can be damaged.

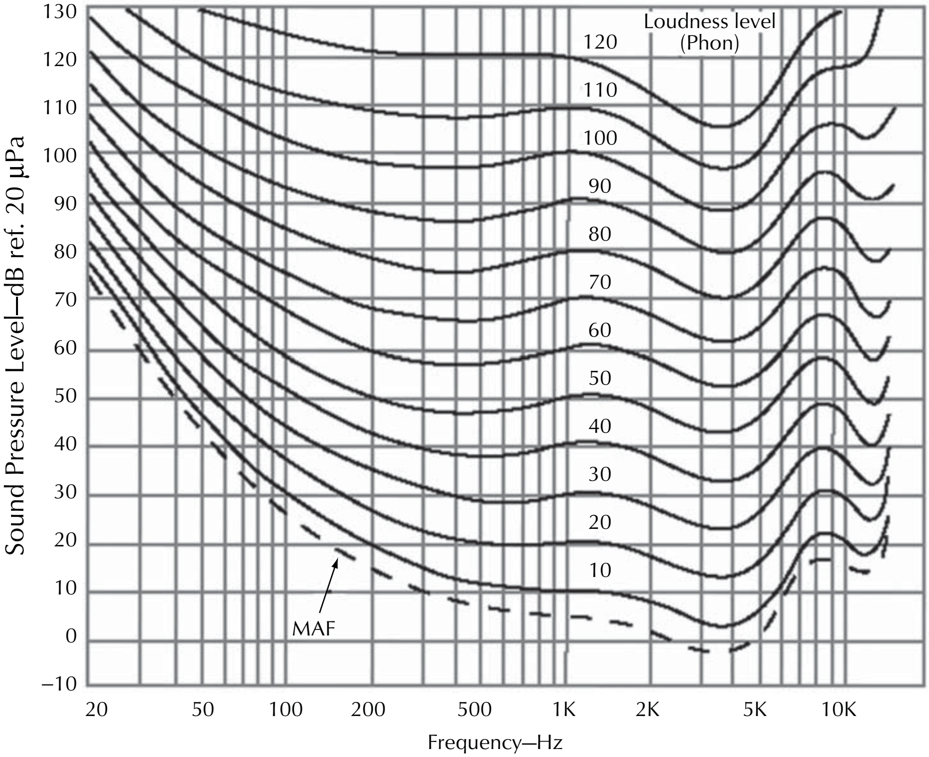


Figure 2-17. The equal-loudness contours. Illustration courtesy Syn-Aud-Con.

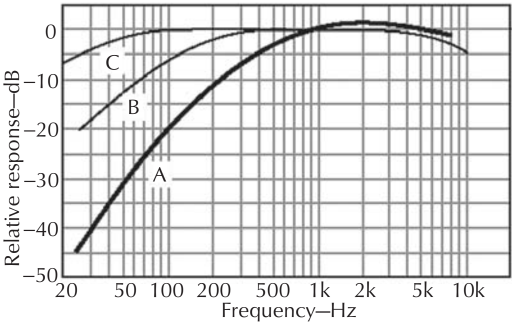


Figure 2-18. Weighting scales for measuring sound levels. Illustration courtesy Syn-Aud-Con.

The Occupational Safety and Health Administration (OSHA) is responsible for assuring that public spaces remain in compliance regarding sound exposure. Sound systems are a major source of high level sounds and should work within OSHA guidelines. Tinnitus, or ringing in the ears, is one symptom of excessive sound exposure.

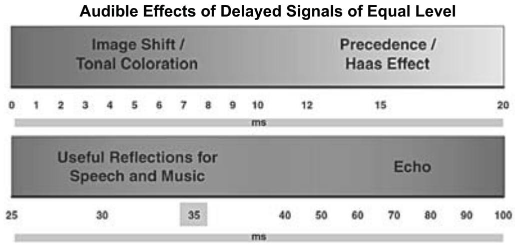

Audible Effects of Delayed Signals of Equal Level
Image Shift /
Tonal Coloration
Precedence /
Haas Effect
0
1
2
3
4
5
6
7
8
9
10
12
15
20
ms
Useful Reflections for
Speech and Music
Echo
25
30
35
40
50
60
70
80
90
100
ms

Figure 2-19. The time offset between sound arrivals will determine if the secondary arrival is useful or harmful in conveying information to the listener. The direction of arrival is also important and is considered by acousticians when designing auditoriums. Courtesy Syn-Aud-Con.

2.10 Monitoring Audio Program Material

The complex nature of the audio waveform necessitates specialized instrumentation for visual monitoring. Typical voltmeters are not suitable for anything but the simplest waveforms, such as sine waves. There are two aspects of the audio signal that are of interest to the system operator. The peaks of the program material must not exceed the peak output capability of any component in the system. Ironically the peaks have little to do with the perceived loudness of the signal or the electrical or acoustic power generated by it. Both of these parameters are more closely tied to the rms (root-mean-square) value of the signal. Measurement of the true rms value of a waveform requires specialized equipment that integrates energy over a time span, much like the hearing system does. This integrated data will better correlate with the perceived loudness of the sound event. So audio practitioners need to monitor at least two aspects of the audio signal—its relative loudness (related to the rms level) and peak levels. Due to the complexity of true rms monitoring, most meters display an average value that is an approximation of the rms value of the program material.

Many audio processors have instrumentation to monitor either peak or average levels, but few can track both simultaneously. Most mixers have a VI (volume indicator) meter that reads in VU (volume units). Such meters are designed with ballistic properties that emulate the human hearing system and are useful for tracking the perceived loudness of the signal. Meters of this type all but ignore the peaks in the program material, making them unable to display the available headroom in the system or clipping in a component. Signal processors usually have a peak LED that responds fast enough to indicate peaks that are at or near the component's clipping point. Many recording systems have PPM (peak program meters) that track the peaks but reveal little about the relative loudness of the waveform.

Fig. 2-20 shows an instrument that monitors both peak and relative loudness of the audio program material. Both values are displayed in relative dB, and the difference between them is the approximate *crest factor* of the program material. Meters of this type yield a more complete picture of the audio event, allowing both loudness and available headroom to be observed simultaneously.

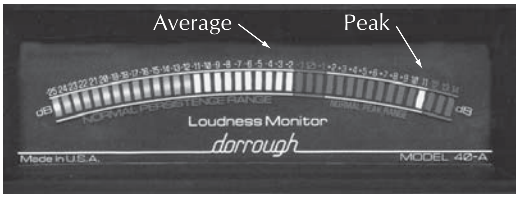


Figure 2-20. A meter that can display both average and peak levels simultaneously. Courtesy Dorrough Electronics.

2.11 Sound Propagation

Sound waves are emitted from acoustic sources—devices that move to modulate the ambient atmospheric pressure. Loudspeakers become intentional acoustic sources when they are driven with waveforms that cause them to vibrate at frequencies within the bandwidth of the human listener. A point source is a device that radiates sound from one point in space. A true point source is an abstract idea and is not physically realizable, as it would be of infinitesimal size. This does not prevent the use of the concept to describe the characteristics of devices that are physically realizable.

Let us consider the properties of some idealized acoustic sources—not ideal in that they would be desirable for sound reinforcement use, but ideal in respect to their behavior as predictable radiators of acoustic energy.

2.11.1 The Point Source

A point source with 100% efficiency would produce 1 watt of acoustical power from one watt of applied electrical power. No heat would result, since all of the electrical power is converted. The energy radiated from the source would travel equally in all directions from the source. Directional energy radiation is accomplished by interfering with the emerging wave. Since interference would require a finite size, a true infinitesimal point source would be omnidirectional. We will introduce the effects of interference later.

Using 1 pW (picowatt) as a power reference, the sound power level produced by 1 acoustic watt will be

$$\begin{aligned} L_W &= 10\log\frac{1\,\mathrm{W}}{10^{-12}\,\mathrm{W}} \\ &= 120\ \mathrm{dB} \end{aligned} \tag{2-20}$$

Note that the sound power is not dependent on the distance from the source. A sound power level of $L_W = 120$ dB would represent the highest continuous sound power level that could result from 1 W of continuous electrical power. All real-world devices will fall short of this ideal, requiring that they be rated for efficiency and power dissipation.

Let us now select an observation point at a distance 0.282 m from the sound source. As the sound energy propagates, it forms a spherical wave front. At 0.282 m this wave front will have a surface area of one square meter. As such, the one watt of radiated sound power is passing through a surface area of 1 m^2.

$$L_I = 10\log\frac{1\,\mathrm{W/m^2}}{10^{-12}\,\mathrm{W/m^2}} \qquad (2\text{-}21)$$
$$= 120\ \mathrm{dB}$$

This is the sound intensity level L_I of the source and represents the amount of power flowing through the surface of a sphere of 1 square meter. Again, this is the highest intensity level that could be achieved by an omnidirectional device of 100% efficiency. L_I can be manipulated by confining the radiated energy to a smaller area. The level benefit gained at a point of observation by doing such is called the *directivity index* (DI) and is expressed in decibels. All loudspeakers suitable for sound reinforcement should exploit the benefits of directivity control.

For the ideal device described, the sound pressure level L_P (or commonly *SPL*) at the surface of the sphere will be numerically the same as the L_W and L_I (L_P = 120 dB) since the sound pressure produced by 1 W will be 20 Pa. This L_P is only for one point on the sphere, but since the source is omnidirectional, all points on the sphere will be the same. To summarize, at a distance of 0.282 m from a point source, the sound power level, sound intensity level, and sound pressure level will be numerically the same. This important relationship is useful for converting between these quantities, Fig. 2-21.

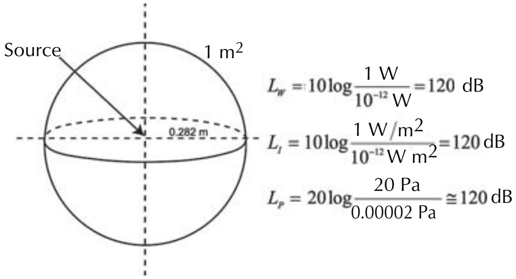

Source
1 m²
0.282 m
$L_W = 10\log\frac{1\ \mathrm{W}}{10^{-12}\ \mathrm{W}} = 120$ dB
$L_I = 10\log\frac{1\ \mathrm{W/m^2}}{10^{-12}\ \mathrm{W\ m^2}} = 120$ dB
$L_P = 20\log\frac{20\ \mathrm{Pa}}{0.00002\ \mathrm{Pa}} \cong 120$ dB

Figure 2-21. This condition forms the basis of the standard terminology and relationships used to describe sound radiation from loudspeakers. Courtesy Syn-Aud-Con.

Let us now consider a point of observation that is twice as far from the source. As the wave continues to spread, its total area at a radius of 0.564 m will be four times the area at 0.282 m. When the sound travels twice as far, it spreads to cover four times the area. In decibels, the sound level change from point one to point two is

$$\Delta L_p = 20\log\frac{0.564}{0.282}$$
$$= 6\ \mathrm{dB}$$

This behavior is known as the inverse-square law (ISL), Fig. 2-22. The ISL describes the level attenuation versus distance for a point source radiator due to the spherical spreading of the emerging waves. Frequency dependent losses will be incurred from atmospheric absorption, but those will not be considered here. Most loudspeakers will roughly follow the inverse square law level change with distance at points remote from the source, Fig. 2-23.

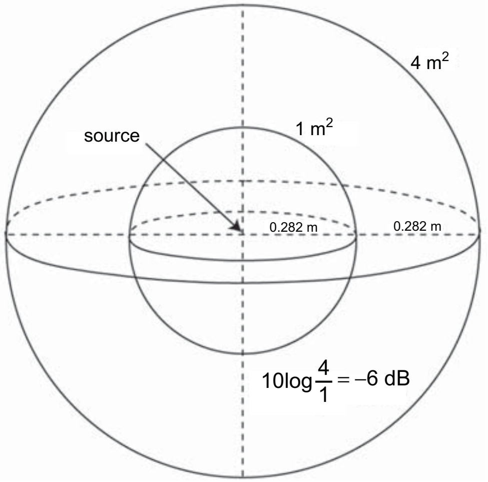


Figure 2-22. When the distance to the source is doubled, the radiated sound energy will be spread over twice the area. Both L_I and L_P will drop by 6 dB. Courtesy Syn-Aud-Con.

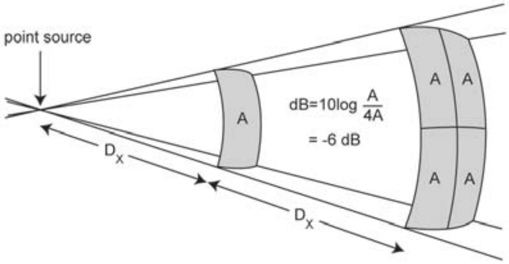


Figure 2-23. The ISL is also true for directional devices in their far field (remote locations from the device). Courtesy Syn-Aud-Con.

2.11.2 The Line Source

Successful sound radiators have been constructed that radiate sound from a line rather than a point. The infinite line source emits a wave that is approximately cylindrical in shape. Since the diverging wave is not expanding in two dimensions, the level change with increasing distance is half that of the point source radiator. The sound level from an ideal line source will decrease at 3 dB per distance doubling rather than 6 dB, Fig. 2-24. It should be pointed out that these relationships are both frequency and line length dependent, and what is being described here is the ideal case. Few commercially available line arrays exhibit this cylindrical behavior over their full bandwidth. Even so, it is useful to allow a mental image of the characteristics of such a device to be formed.

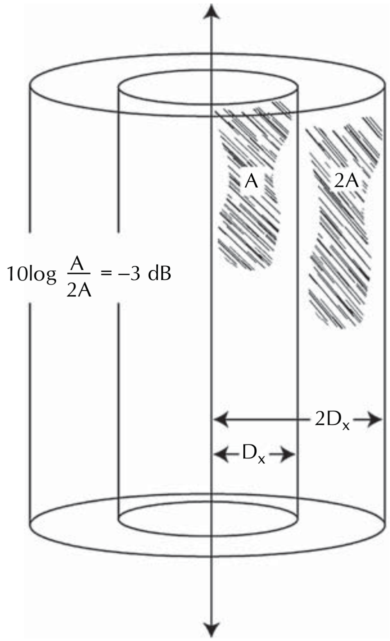


Figure 2-24. Line sources radiate a cylindrical wave (ideal case). The level drop versus distance is less than for a point source. Courtesy Syn-Aud-Con.

If the line source is finite in length (as all real-world sources will be), then there will be a phase differential between the sound radiated from different points on the source to a specific point in space. All of the points will be the most in phase on a plane perpendicular from the array and equidistant from the end points of the array. As the point of observation moves away from the midpoint, phase interaction will produce lobes in the radiated energy pattern. The lobes can be suppressed by clever design, allowing the wave front to be confined to a very narrow vertical angle, yet with wide horizontal coverage. Such a radiation pattern is ideal for some applications, such as a broad, flat audience plane that must be covered from ear height. Digital signal processing has produced well-behaved line arrays that can project sound to great distances. Some incorporate an adjustable delay for each element to allow steering of the radiation lobe. Useful designs for auditoriums are at least 2 meters in vertical length.

While it is possible to construct a continuous line source using ribbon drivers, etc., most commercially available designs are made up of closely spaced discrete loudspeakers or loudspeaker systems and are more properly referred to as line arrays, Fig. 2-25.

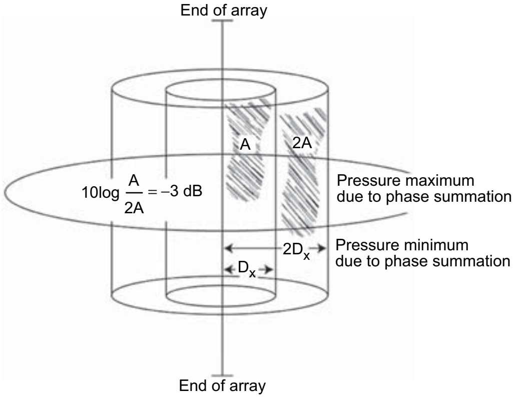

End of array
A
2A
$10\log\frac{A}{2A} = -3$ dB
Pressure maximum due to phase summation
$2D_x$
D_x
Pressure minimum due to phase summation
End of array

Figure 2-25. The finite line array has gained wide acceptance among system designers, allowing wide audience coverage with minimal energy radiation to room surfaces. Courtesy Syn-Aud-Con.

2.12 Conclusion

The material in this chapter was carefully selected to expose the reader to a broad spectrum of principles regarding sound reinforcement systems. As a colleague once put it, "Sound theory is like an onion. Every time you peel off a layer another lies beneath it!" Each of these topics can be taken to higher levels, and many have been by other authors within this textbook. The reader is encouraged to use this information as a springboard into a life-long study of audio and acoustics. We are called upon to spend much of our time learning about new technologies. It must be remembered that new methods come from the mature body of principles and practices that have been handed down by those who came before us. Looking backward can have some huge rewards.

If I can see farther than those who came before me, it is because I am standing on their shoulders.

Sir Isaac Newton

Bibliography

D. Davis and C. Davis, *Sound System Engineering*, Boston: Focal Press, 1997.
Glen Ballou, *Handbook for Sound Engineers, 2nd Edition*, Boston: Focal Press, 1991.
F. Alton Everest, *Master Handbook of Acoustics, 3rd Edition*, TAB Books, 1998.
Leo Baranek, *Acoustics*, New York: McGraw-Hill, 1954.
Harry Olson, *Acoustical Engineering*, Professional Audio Journals, Inc., Philadelphia, PA.

This page intentionally left blank

Chapter 3

by Peter Xinya Zhang

3.1 Psychoacoustics and Subjective Quantities 43
3.2 Ear Anatomy and Function 43
  3.2.1 Pinna 43
  3.2.2 Temporal Bones 44
  3.2.3 Ear Canal 45
  3.2.4 Middle Ear 45
    3.2.4.1 Acoustic Reflex 45
  3.2.5 Inner Ear 45
3.3 Frequency Selectivity 47
  3.3.1 Frequency Tuning 47
  3.3.2 Masking and Its Application in Audio Encoding 48
  3.3.3 Auditory Filters and Critical Bands 49
3.4 Nonlinearity of the Ear 49
3.5 Perception of Phase 50
3.6 Auditory Area and Thresholds 50
3.7 Hearing Over Time 51
3.8 Loudness 51
  3.8.1 Equal Loudness Contours and Loudness Level 51
  3.8.2 Level Measurements with A-, B-, and C-Weightings 52
  3.8.3 Loudness in Sones 52
  3.8.4 Loudness versus Bandwidth 53
  3.8.5 Loudness of Impulses 53
  3.8.6 Perception of Dynamic Changes 54
3.9 Pitch 54
  3.9.1 The Unit of Pitch 55
  3.9.2 Perception of Pure and Complex Tones 55
  3.9.3 Phase-Locking and the Range of Pitch Sensation 56
  3.9.4 Frequency Difference Limen 56
  3.9.5 Dependence of Pitch on Level 56
  3.9.6 Perfect Pitch 56
  3.9.7 Other Pitch Effects 57
3.10 Timbre 57
3.11 Binaural and Spatial Hearing 57
  3.11.1 Localization Cues for Left and Right 57
  3.11.2 Localization on Sagittal Planes 58
  3.11.3 Externalization 59
  3.11.4 Precedence (Hass) Effect 60
  3.11.5 Franssen Effect 60
  3.11.6 Cocktail Party Effect and Improvement of Signal Detection 60
  3.11.7 Distance Perception 60
Further Reading 61
References 61

This page intentionally left blank

3.1 Psychoacoustics and Subjective Quantities

Unlike other senses, it is surprising how limited our vocabulary is when talking about hearing.[1] Especially in the audio industry, we do not often discriminate between subjective and objective quantities. For instance, the quantities of frequency, level, spectrum, etc. are all objective, in a sense that they can be measured with a meter or an electronic device; whereas the concepts of pitch, loudness, timbre, etc. are subjective, and they are auditory perceptions in our heads. Psychoacoustics investigates these subjective quantities (i.e., our perception of hearing), and their relationship with the objective quantities in acoustics. *Psychoacoustics* got its name from a field within psychology—i.e., recognition science—which deals with all kinds of human perceptions, and it is an interdisciplinary field of many areas, including psychology, acoustics, electronic engineering, physics, biology, physiology, computer science, etc.

Although there are clear and strong relationships between certain subjective and objective quantities—e.g., pitch versus frequency—other objective quantities also have influences. For example, changes in sound level can affect pitch perception. Furthermore, because no two persons are identical, when dealing with perceptions as in psychoacoustics, there are large individual differences, which can be critical in areas such as sound localization.[2] In psychoacoustics, researchers have to consider both average performances among population and individual variations. Therefore, psychophysical experiments and statistical methods are widely used in this field.

Compared with other fields in acoustics, psychoacoustics is relatively new, and has been developing greatly. Although many of the effects have been known for some time (e.g., Hass effect[3]), new discoveries have been found continuously. To account for these effects, models have been proposed. New experimental findings might invalidate or modify old models or make certain models more or less popular. This process is just one representation of how we develop our knowledge. For the purpose of this handbook, we will focus on summarizing the known psychoacoustic effects rather than discussing the developing models.

3.2 Ear Anatomy and Function

Before discussing various psychoacoustic effects, it is necessary to introduce the physiological bases of those effects, namely the structure and function of our auditory system. The human ear is commonly considered in three parts: the outer ear, the middle ear, and the inner ear. The sound is gathered (and as we shall see later, modified) by the external ear called the *pinna* and directed down the *ear canal* (auditory meatus). This canal is terminated by the *tympanic membrane* (eardrum). These parts constitute the outer ear, as shown in Figs. 3-1 and 3-2. The other side of the eardrum faces the middle ear. The middle ear is air filled, and pressure equalization takes place through the *eustachian tube* opening into the *pharynx* so normal atmospheric pressure is maintained on both sides of the eardrum. Fastened to the eardrum is one of the three ossicles, the *malleus* which, in turn, is connected to the *incus* and *stapes*. Through the rocking action of these three tiny bones the vibrations of the eardrum are transmitted to the *oval window* of the *cochlea* with admirable efficiency. The sound pressure in the liquid of the cochlea is increased some 30–40 dB over the air pressure acting on the eardrum through the mechanical action of this remarkable middle ear system. The clear liquid filling the cochlea is incompressible, like water. The *round window* is a relatively flexible pressure release allowing sound energy to be transmitted to the fluid of the cochlea via the oval window. In the inner ear the traveling waves set up on the basilar membrane by vibrations of the oval window stimulate *hair cells* that send nerve impulses to the brain.

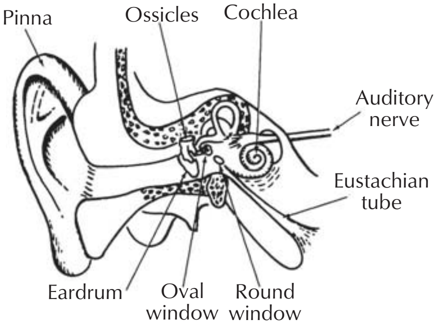


Figure 3-1. A cross-section of the human ear showing the relationship of the various parts.

3.2.1 Pinna

The *pinna*, or the human *auricle*, is the most lateral (i.e., outside) portion of our auditory system. The beauty of these flaps on either side of our head may be questioned, but not the importance of the acoustical function they serve. Fig. 3-3 shows an illustration of various parts of the human pinna. The entrance to the ear canal, or *concha*, is most important acoustically for filtering because it contains the largest air volume in a pinna. Let

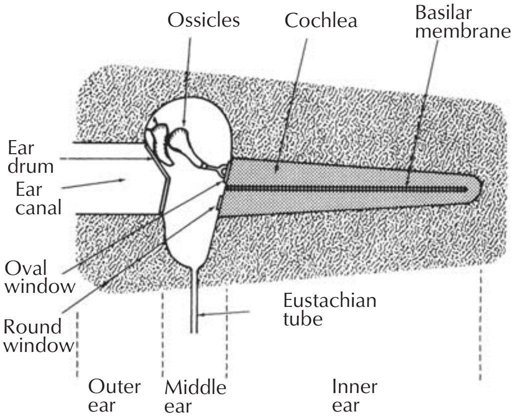


Figure 3-2. Highly idealized portrayal of the outer ear, middle ear, and inner ear.

us assume for the moment that we have no pinnae, just holes in the head, which is actually a simplest model for human hearing, called the *spherical head model*. Cupping our hands around the holes would make sounds louder as more sound energy is directed into the opening. How much does the pinna help in directing sound energy into the ear canal? We can get some idea of this by measuring the sound pressure at the opening of the ear canal with and without the hand behind the ear. Wiener and Ross[4] did this and found a gain of 3 to 5 dB at most frequencies, but a peak of about 20 dB in the vicinity of 1500 Hz. Fig. 3-4 shows the transfer function measured by Shaw,[5] and the curves numbered 3 and 4 are for concha and pinna flange, respectively. The irregular and asymmetric shape of a pinna is not just for aesthetic reasons. In Section 3.11, we will see that it is actually important for our ability to localize sounds and to aid in spatial-filtering of unwanted conversations.

3.2.2 Temporal Bones

On each of the left and right sides of our skull, behind the pinna, there is a thin, fanlike bone—namely, the *temporal bone*—covering the entire human ear, except for the pinna. This bone can be further divided into four portions—i.e., the *squamous*, *mastoid*, *tympanic* and *petrous portions*. The obvious function for the temporal bone is to protect our auditory system. Other than cochlear implant patients, whose temporal bone has to be partly removed during a surgery, people might not pay much attention to it, especially regarding acoustics. However the sound energy that propagates through the bone into our inner ear, as opposed to through the ear canal and middle ear, is actually fairly significant. For

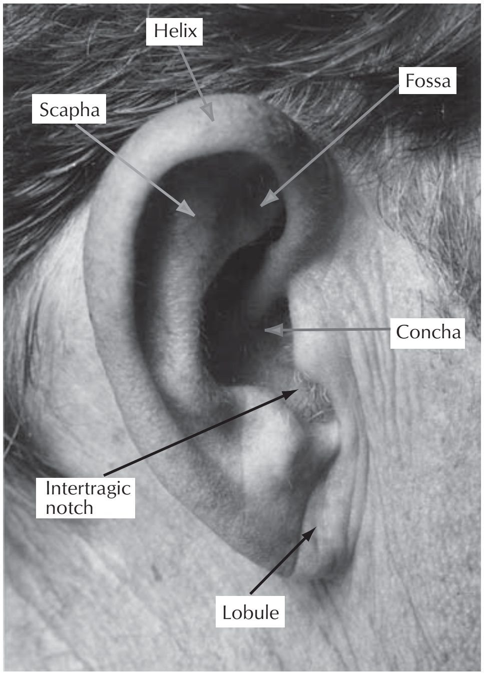


Figure 3-3. The human outer ear, the pinna, with identification of some of the folds, cavities, and ridges that have significant acoustical effect.

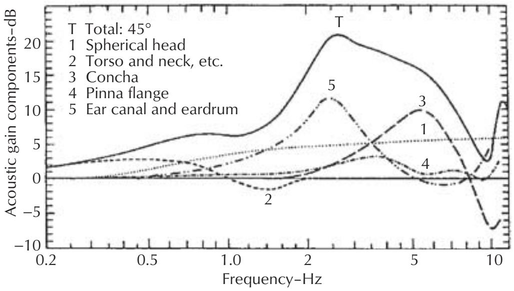


Figure 3-4. The average pressure gain contributed by the different components of the outer ear in humans. The sound source is in the horizontal plane, 45° from straight ahead. (After Shaw, Reference 5.)

patients with conductive hearing loss—e.g., damage of middle ear—there are currently commercially available devices, which look something like headphones and are placed on the temporal bone. People with normal hearing can test it by plugging their ears while wearing the device. Although the timbres sound quite different from normal hearing, the filtered speech is clear enough to understand. Also because of this bone conduction, along with other effects such as acoustic reflex, which will be discussed in Section 3.2.4.1, one hears his or her own voice differently from how other people hear the voice. While not receiving much attention in everyday life, it might be sometimes very important. For example, an experienced voice teacher often asks a student singer to record his or her own singing and playback with an audio system. The recording will sound unnatural to the singer but will be a more accurate representation of what the audience hears.

3.2.3 Ear Canal

The ear canal has a diameter about 5 to 9 mm and is about 2.5 cm long. It is open to the outside environment at the concha, and is closed at the tympanic membrane. Acoustically, it can be considered as a closed pipe whose cross-sectional shape and area vary along its length. Although being bended and irregular in shape, the ear canal does demonstrate the modal characteristic of a closed pipe. It has a fundamental frequency of about 3 kHz, corresponding to a quarter wavelength close to the length of the ear canal. Because of this resonant frequency, our hearing is most sensitive to a frequency band around 3 kHz, which is, not just by coincidence, the most important frequency band of human speech. On Fig. 3-4, the number 5 curve shows the effect of the ear canal, taking the eardrum into account as well. As can be seen, there is an approximately 11 dB of gain at around 2.5 kHz. After combining all the effects of head, torso and neck, pinna, ear canal and eardrum, the total transfer function is the curve marked with a letter T on Fig. 3-4. It is relatively broadly tuned between 2 and 7 kHz, with as much as 20 dB of gain. Unfortunately, because of this resonance, in very loud noisy environments with broadband sound, hearing damage usually first happens around 4 kHz.

3.2.4 Middle Ear

The outer ear, including the pinna and the ear canal, ends at the eardrum. It is an air environment with low impedance. On the other hand, the inner ear, where the sensory cells are, is a fluid environment with high impedance. When sound (or any wave) travels from one medium to another, if the impedances of the two media do not match, much of the energy would be reflected at the surface, without propagating into the second medium. For the same reason, we use microphones to record in the air and hydrophones to record under water. To make our auditory system efficient, the most important function of the middle ear is to match the impedances of outer and inner ears. Without the middle ear, we would suffer a hearing loss of about 30 dB (by mechanical analysis[6] and experiments on cats[7]).

A healthy middle ear (without middle ear infection) is an air-filled space. When swallowing, the eustachian tube is open to balance the air pressure inside the middle ear and that of the outside world. Most of the time, however, the middle ear is sealed from the outside environment. The main components of the middle ear are the three ossicles, which are the smallest bones in our body: the malleus, incus, and stapes. These ossicles form an ossiclar chain, which is firmly fixed on the eardrum and the oval window on each side. Through mostly three types of mechanical motions—namely piston motion, lever motion and buckling motion[8]—the acoustic energy is transferred into the inner ear effectively. The middle ear can be damaged temporarily by middle ear infection, or permanently by genetic disease. Fortunately, with current technology, doctors can rebuild the ossicles with titanium, the result being a total recovering of hearing.[9] Alternatively one can use devices that rely on bone conduction.[10]

3.2.4.1 Acoustic Reflex

There are two muscles in the middle ear: the *tensor tympani* that is attached to the malleus, and the *stapedius muscle* that is attached to the stapes. Unlike other muscles in our bodies, these muscles form an angle with respect to the bone, instead of along the bone, which makes them very ineffective for motion. Actually the function of these muscles is for changing the stiffness of the ossicular chain. When we hear a very loud sound—i.e., at least 75 dB higher than the hearing threshold—when we talk or sing, when the head is touched, or when the body moves,[11] these middle ear muscles will contract to increase the stiffness of the ossicular chain, which makes it less effective, so that our inner ear is protected from exposure to the loud sound. However, because this process involves a higher stage of signal processing, and because of the filtering features, this protection works only for slow onset and low-frequency sound (up to 1.2 kHz[12]) and is not effective for noises such as an impulse or noise with high frequencies (e.g., most of the music recordings today).

3.2.5 Inner Ear

The inner ear, or the *labyrinth*, is composed of two systems: the *vestibular system,* which is critical to our sense of balance, and the auditory system which is used for hearing. The two systems share fluid, which is separated from the air-filled space in the middle ear by the oval window and the round window. The auditory portion of the inner ear is the snail-shaped cochlea. It is a mechanical-to-electrical transducer and a frequency-selective analyzer, sending coded nerve impulses to the brain. This is represented crudely in Fig. 3-5. A rough sketch of the cross section of the cochlea is shown in Fig. 3-6. The cochlea, throughout its length (about 35 mm if stretched out straight), is divided by *Reissner's membrane* and the *basilar membrane* into three separate compartments—namely, the *scala vestibuli*, the *scala media,* and the *scala tympani*. The *scala vestibuli* and the *scala tympani* share the same fluid, *perilymph*, through a small hole, the *helicotrema*, at the apex; while the *scala media* contains another fluid, *endolymph*, which contains higher density of potassium ions facilitating the function of the hair cells. The *basilar membrane* supports the *Organ of Corti*, which contains the hair cells that convert the relative motion between the *basilar membrane* and the *tectorial membrane* into nerve pulses to the auditory nerve.

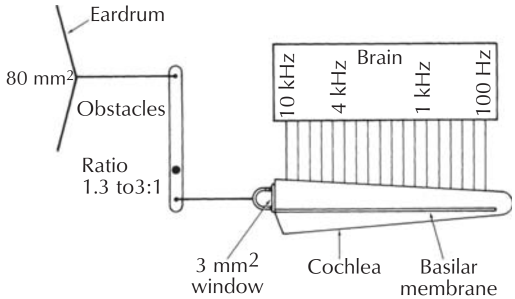


Figure 3-5. The mechanical system of the middle ear.

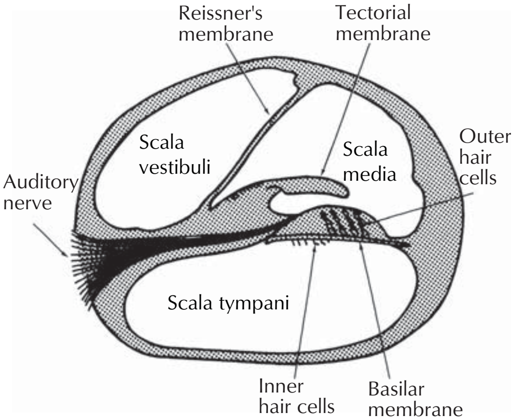


Figure 3-6. Cross-sectional sketch of the cochlea.

When an incident sound arrives at the inner ear, the vibration of the stapes is transported into the scala vestibuli through the oval window. Because the cochlear fluid is incompressible, the round window connected to the scala tympani vibrates accordingly. Thus, the vibration starts from the base of the cochlea, travels along the scala vestimbuli, all the way to the apex, and then through the helicotrema into the scala tympani, back to the base, and eventually ends at the round window. This establishes a traveling wave on the basilar membrane for frequency analysis. Each location at the basilar membrane is most sensitive to a particular frequency—i.e., the characteristic frequency—although it also responds to a relatively broad frequency band at smaller amplitude. The basilar membrane is narrower (0.04 mm) and stiffer near the base, and wider (0.5 mm) and looser near the apex. (By contrast, when observed from outside, the cochlea is wider at the base and smaller at the apex.) Therefore, the characteristic frequency decreases gradually and monotonically from the base to the apex, as indicated in Fig. 3-5. The traveling-wave phenomenon illustrated in Figs. 3-7 and 3-8 shows the vibration patterns—i.e., amplitude versus location—for incident pure tones of different frequencies. An interesting point in Fig. 3-8 is that the vibration pattern is asymmetric, with a slow tail close to the base (for high frequencies) and a steep edge close to the apex (for low frequencies). Because of this asymmetry, it is easier for the low frequencies to mask the high frequencies than vice versa.

Within the Organ of Corti on the basilar membrane, there are a row of *inner hair cells* (IHC), and three to five rows of *outer hair cells* (OHC), depending on location. There are about 1500 *IHC*s and about 3500 OHCs. Each hair cell contains stereociliae (hairs) that vibrate corresponding to the mechanical vibration in the fluid around them. Because each location on the basilar membrane is most sensitive to its own characteristic frequency, the hair cells at the location also respond most to its characteristic frequency. The IHCs are sensory cells, like microphones, which convert mechanical vibration into electrical signal—i.e., neural firings. The OHCs, on the other hand, change their shapes according to the control signal received from efferent nerves. Their function is to give an extra gain or attenuation, so that the output of the IHC is tuned to the characteristic frequency much more sharply than the IHC itself. Fig. 3-9 shows the tuning curve (output level vs. frequency) for a particular location on the basilar membrane with and without functioning OHCs. The tuning curve is much broader with poor frequency selectivity when the OHCs do not function. The OHCs make our auditory system an active device, instead of a passive microphone. Because the OHCs are active and

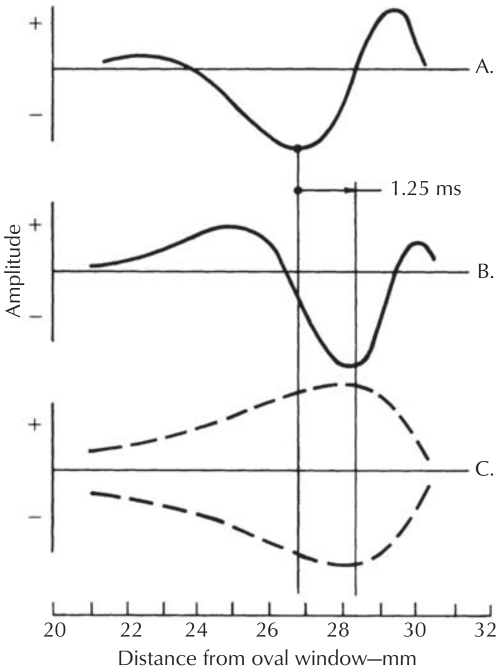


Figure 3-7. A traveling wave on the basilar membrane of the inner ear. (After von Békésy, Reference 13.) The exaggerated amplitude of the basilar membrane for a 200 Hz wave traveling from left to right is shown at A. The same wave 1.25 ms later is shown at B. These traveling 200 Hz waves all fall within the envelope at C.

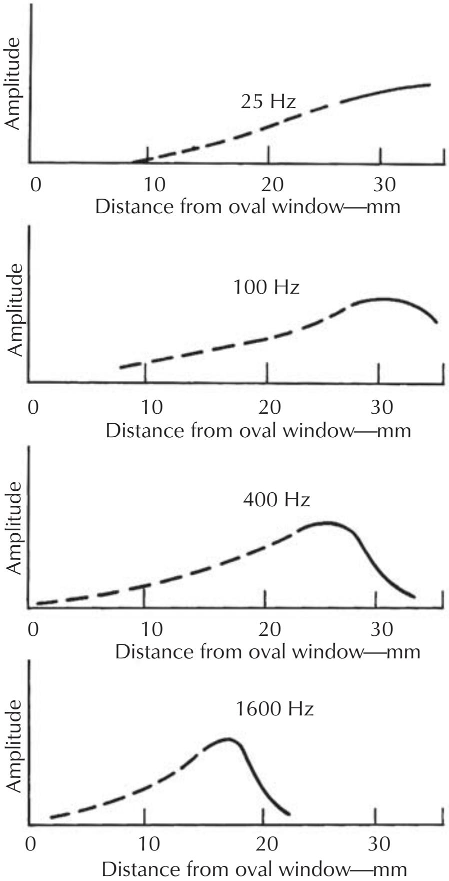


Figure 3-8. An illustration of vibration patterns of the hair cells on the basilar membrane for various incident pure tones. There is a localized peak response for each audible frequency. (After von Békésy, Reference 13.)

consume a lot of energy and nutrition, they are usually damaged first due to loud sound or ototoxic medicines (i.e., medicine that is harmful to the auditory system). Not only does this kind of hearing loss make our hearing less sensitive, it also makes our hearing less sharp. Thus, as is easily confirmed with hearing-loss patients, simply adding an extra gain with hearing aids would not totally solve the problem.

3.3 Frequency Selectivity

3.3.1 Frequency Tuning

As discussed in Section 3.2.5, the inner hair cells are sharply tuned to the characteristic frequencies with help from the outer hair cells. This tuning character is also conserved by the auditory neurons connecting to the inner hair cells. However, this tuning feature varies with level. Fig. 3-10 shows a characteristic diagram of tuning curves from a particular location on the basilar membrane at various levels. As can be seen in this graph, as level increases, the tuning curve becomes broader, indicating less frequency selectivity. Thus, in order to hear music more sharply, one should play back at a relatively low level. Moreover, above 60 dB, as level increases, the characteristic frequency decreases. Therefore when one hears a tone at a high level, a neuron that is normally tuned at a higher characteristic frequency is now best tuned to the tone. Because eventually the brain perceives pitch based on neuron input, at high levels, without knowing that the characteristic frequency has decreased, the brain hears the pitch to be sharp.

Armed with this knowledge, one would think that someone who was engaged in critical listening—a recording engineer, for example—would choose to listen at moderate to low levels. Why then do so many

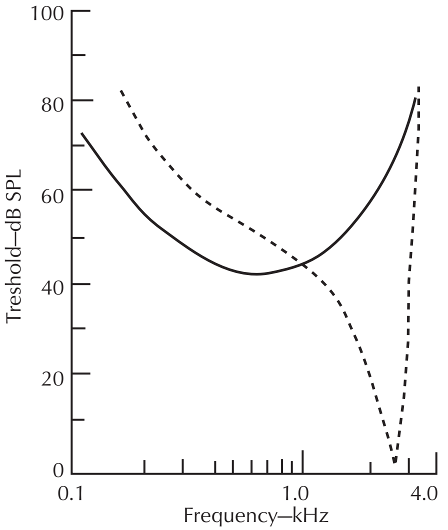


Figure 3-9. Tuning curve with (solid) and without (dashed) functioning outer hair cells. (Liberman and Dodds, Reference 14.)

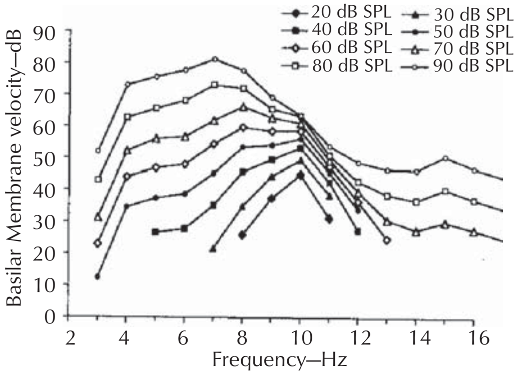


Figure 3-10. Tuning curve at various levels at a particular location of the basilar membrane of a chinchilla. (Plack, Reference 15, p90, Ruggero et al., Reference 16.)

audio professionals choose to monitor at very high levels? There could be many reasons. Loud levels may be more exciting. It may simply be a matter of habit. For instance, an audio engineer normally turns the volume to his or her customary level fairly accurately. Moreover, because frequency selectivity is different at different levels, an audio engineer might choose to make a recording while listening at a "realistic" or "performance" level rather than monitoring at a level that is demonstratedly more accurate. Finally, of course, there are some audio professionals who have lost some hearing already, and in order to pick up certain frequency bands they keep on boosting the level, which unfortunately further damages their hearing.

3.3.2 Masking and Its Application in Audio Encoding

Suppose a listener can barely hear a given acoustical signal under quiet conditions. When the signal is playing in presence of another sound (called "a *masker*"), the signal usually has to be stronger so that the listener can hear it.[17] The masker does not have to include the frequency components of the original signal for the masking effect to take place, and a masked signal can already be heard when it is still weaker than the masker.[18]

Masking can happen when a signal and a masker are played simultaneously (simultaneous masking), but it can also happen when a masker starts and ends before a signal is played. This is known as forward masking. Although it is hard to believe, masking can also happen when a masker starts *after* a signal stops playing! In general, the effect of this backward masking is much weaker than forward masking. Forward masking can happen even when the signal starts more than 100 ms after the masker stops,[19] but backward masking disappears when the masker starts 20 ms after the signal.[20]

The masking effect has been widely used in psychoacoustical research. For example, Fig. 3-10 shows the tuning curve for a chinchilla. For safety reasons, performing such experiments on human subjects is not permitted. However, with masking effect, one can vary the level of a masker, measure the threshold (i.e., the minimum sound that the listener can hear), and create a diagram of a psychophysical tuning curve that reveals similar features.

Besides scientific research, masking effects are also widely used in areas such as audio encoding. Now, with distribution of digital recordings, it is desirable to reduce the sizes of audio files. There are *lossless encoders*, which is an algorithm to encode the audio file into a smaller file that can be completely reconstructed with another algorithm (decoder). However, the file sizes of the lossless encoders are still relatively large. To further reduce the size, some less important information has to be eliminated. For example, one might eliminate high frequencies, which is not too bad for speech communication. However, for music, some important quality might be lost. Fortunately, because of the masking effect, one can eliminate some weak sounds that are masked so that listeners hardly notice the difference. This technique has been widely used in audio encoders, such as MP3.

3.3.3 Auditory Filters and Critical Bands

Experiments show that our ability to detect a signal depends on the bandwidth of the signal. Fletcher (1940)[18] found that, when playing a tone in the presence of a bandpass masker, as the masker bandwidth was increased while keeping the overall level of the masker unchanged, the threshold increased as bandwidth increased up to a certain limit, beyond which the threshold remained constant. One can easily confirm that, when listening to a bandpass noise with broadening bandwidth and constant overall level, the loudness is unchanged, until a certain bandwidth is reached, and beyond that bandwidth the loudness increases as bandwidth increases, although the reading of an SPL meter is constant. An explanation to account for these effects is the concept of auditory filters. Fletcher proposed that, instead of directly listening to each hair cell, we hear through a set of auditory filters, whose center frequencies can vary or overlap, and whose bandwidth is varying according to the center frequency. These bands are referred to as *critical bands* (CB). Since then, the shape and bandwidth of the auditory filters have been carefully studied. Because the shape of the auditory filters is not simply rectangular, it is more convenient to use the *equivalent rectangular bandwidth* (ERB), which is the bandwidth of a rectangular filter that gives the same transmission power as the actual auditory filter. Recent study by Glasberg and Moore (1990) gives a formula for ERB for young listeners with normal hearing under moderate sound pressure levels[21]:

$$ERB = 24.7(4.37F + 1)$$

where,
the center frequency of the filter F is in kHz,
ERB is in Hz.

Sometimes, it is more convenient to use an ERB number as in Eq. 3-1,[21] similar to the Bark scale proposed by Zwicker et al.[22]:

$$ERB\ Number = 21.4\log(4.37F + 1) \tag{3-1}$$

where,
the center frequency of the filter F is in kHz.

Table 3-1 shows the ERB and Bark scale as a function of the center frequency of the auditory filter. The Bark scale is also listed as a percentage of center frequency, which can then be compared to filters commonly used in acoustical measurements: octave (70.7%), half octave (34.8%), one-third octave (23.2%), and one-sixth octave (11.6%) filters. The ERB is shown in Fig. 3-11 as a function of frequency. One-third octave filters which are popular in audio and have been widely used in acoustical measurements ultimately have their roots in the study of human auditory response. However, as Fig. 3-11 shows, the ERB is wider than ⅓ octave for frequencies below 200 Hz; is smaller than ⅓ octave for frequencies above 200 Hz; and, above 1 kHz, it approaches 1/6 octave.

Table 3-1. Critical Bandwidths of the Human Ear

Critical Band No	Center Frequency Hz	Bark Scale		Equivalent Rectangular Band (ERB), Hz
		(Hz)	%	
1	50	100	200	33
2	150	100	67	43
3	250	100	40	52
4	350	100	29	62
5	450	110	24	72
6	570	120	21	84
7	700	140	20	97
8	840	150	18	111
9	1000	160	16	130
10	1170	190	16	150
11	1370	210	15	170
12	1600	240	15	200
13	1850	280	15	220
14	2150	320	15	260
15	2500	380	15	300
16	2900	450	16	350
17	3400	550	16	420
18	4000	700	18	500
19	4800	900	19	620
20	5800	1100	19	780
21	7000	1300	19	990
22	8500	1800	21	1300
23	10,500	2500	24	1700
24	13,500	3500	26	2400

3.4 Nonlinearity of the Ear

When a set of frequencies are input into a linear system, the output will contain only the same set of frequencies, although the relative amplitudes and phases can be adjusted due to filtering. However, for a nonlinear system, the output will include new frequencies that are not present in the input. Because our auditory system has developed mechanisms such as acoustic reflex in the middle ear and the active processes in the inner ear, it is nonlinear. There are two types of nonlinear-

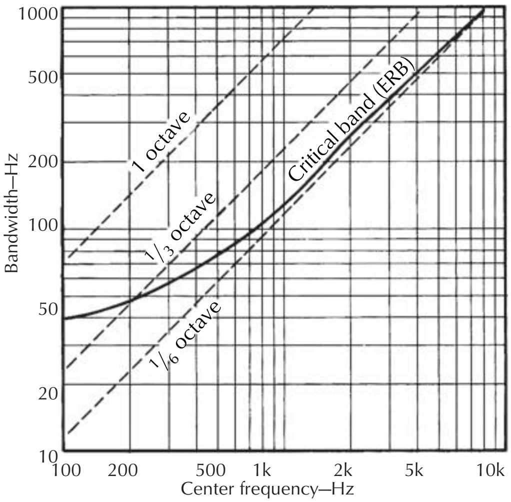


Figure 3-11. A plot of critical bandwidths (calculated *ERB*s) of the human auditory system compared to constant percentage bandwidths of filter sets commonly used in acoustical measurements.

ity—namely, *harmonic distortion* and *combination tones*. Harmonic distortion can be easily achieved by simply distorting a sine-tone. The added new components are harmonics of the original signal. A combination tone happens when there are at least two frequencies in the input. The output might include combination tones according to

$$f_c = |n \times f_1 \pm m \times f_2| \tag{3-2}$$

where,

f_c is the frequency of a combination tone,

f_1 and f_2 are the two input frequencies, and n and m are any integer numbers.

For example, when two tones at 600 and 700 Hz are input, the output might have frequencies such as 100 Hz (= 700 – 600 Hz), 500 Hz (= 2 × 600 – 700 Hz), and 400 Hz (= 3 × 600 – 2 × 700 Hz), etc.

Because the harmonic distortion does not change the perception of pitch, it would not be surprising if we are less tolerant of the combination tones.

Furthermore, because the auditory system is active, even in a completely quiet environment, the inner ear might generate tones. These *otoacoustic emissions*[23] are a sign of a healthy and functioning inner ear, and quite different from the *tinnitus* resulting from exposure to dangerously high sound pressure levels.

3.5 Perception of Phase

The complete description of a given sound includes both an amplitude spectrum and a phase spectrum. People normally pay a lot of attention to the amplitude spectrum, while caring less for the phase spectrum. Yet academic researchers, hi-fi enthusiasts, and audio engineers all have asked, "Is the ear able to detect phase differences?" About the middle of the last century, G. S. Ohm wrote, "Aural perception depends only on the amplitude spectrum of a sound and is independent of the phase angles of the various components contained in the spectrum." Many apparent confirmations of Ohm's law of acoustics have later been traced to crude measuring techniques and equipment.

Actually, the phase spectrum sometimes can be very important for the perception of timbre. For example, an impulse and white noise sound quite different, but they have identical amplitude spectrum. The only difference occurs in the phase spectrum. Another common example is speech: if one scrambles the relative phases in the spectrum of a speech signal, it will not be intelligible. Now, with experimental evidence, we can confirm that our ear is capable of detecting phase information. For example, the neural firing of the auditory nerve happens at a certain phase, which is called the *phase-locking*, up to about 5 kHz.[24] The phase-locking is important for pitch perception. In the brainstem, the information from left and right ears is integrated, and the *interaural phase difference* can be detected, which is important for spatial hearing. These phenomena will be discussed in more detail in Sections 3.9 and 3.11.

3.6 Auditory Area and Thresholds

The auditory area depicted in Fig. 3-12 describes, in a technical sense, the limits of our aural perception. This area is bounded at low sound levels by our threshold of hearing. The softest sounds that can be heard fall on the threshold of hearing curve. Above this line the air molecule movement is sufficient to elicit a response. If, at any given frequency, the sound pressure level is increased sufficiently, a point is reached at which a tickling sensation is felt in the ears. If the level is increased substantially above this threshold of feeling, it becomes painful. These are the lower and upper boundaries of the auditory area. There are also frequency limitations below about 20 Hz and above about 16 kHz, limitations that (like the two thresholds) vary considerably from individual to individual. We are less concerned here about specific numbers than we are about principles. On the auditory area of Fig. 3-12, all the sounds of life are played out—low frequency or high, very soft or very intense. Speech does not utilize the entire auditory area. Its dynamic range and frequency range are quite limited. Music has both a greater dynamic range than speech and a greater frequency range. But even music does not utilize the entire auditory area.

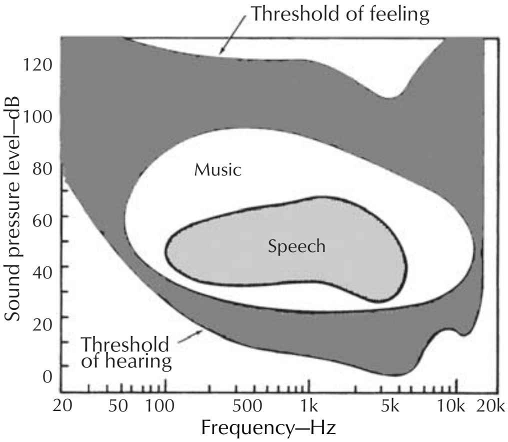


Figure 3-12. All sounds perceived by humans of average hearing acuity fall within the auditory area. This area is defined by the threshold of hearing and the threshold of feeling (pain) and by the low and high frequency limits of hearing. Music and speech do not utilize the entire auditory area available, but music has the greater dynamic range (vertical) and frequency demands (horizontal).

3.7 Hearing Over Time

If our ear was like an ideal Fourier analyzer, in order to translate a waveform into a spectrum, the ear would have to integrate over the entire time domain, which is not practical and, of course, not the case. Actually, our ear only integrates over a limited time window (i.e., a filter on the time axis), and thus we can hear changes of pitch, timbre, and dynamics over time, which can be shown on a *spectrogram* instead of a simple spectrum. Mathematically, it is a wavelet analysis instead of a Fourier analysis. Experiments on gap detection between tones at different frequencies indicate that our temporal resolution is on the order of 100 ms,[25] which is a good estimate of the time window of our auditory system. For many perspectives (e.g., perceptions on loudness, pitch, timbre), our auditory system integrates acoustical information within this time window.

3.8 Loudness

Unlike level or intensity, which are physical or objective quantities, loudness is a listener's subjective perception. As the example in Section 3.3, even if the SPL meter reads the same level, a sound with a wider bandwidth might sound much louder than a sound with a smaller bandwidth. Even for a pure tone, although loudness follows somewhat with level, it is actually a quite complicated function, depending on frequency. A tone at 40 dB SPL is not necessarily twice as loud as another sound at 20 dB SPL. Furthermore, loudness also varies among listeners. For example, a listener who has lost some sensitivity in a certain critical band will perceive any signal in that band to be at a lower level relative to someone with normal hearing.

Although there is no meter to directly measure a subjective quantity such as loudness, psycho-physical scaling can be used to investigate loudness across subjects. Subjects can be given matching tasks, where they are asked to adjust the level of signals until they match, or comparative tasks, where they are asked to compare two signals and estimate the scales for loudness.

3.8.1 Equal Loudness Contours and Loudness Level

By conducting experiments using pure tones with a large population, Fletcher and Munson at Bell Labs (1933) derived equal loudness contours, also known as the *Fletcher-Munson curves*. Fig. 3-13 shows the equal loudness contours later refined by Robinson and Dadson, which have been recognized as an international standard. On the figure, the points on each curve correspond to pure tones, giving the same loudness to an average listener. For example, a pure tone at 50 Hz at 60 dB SPL is on the same curve as a tone at 1 kHz at 30 dB. This means that these two tones have identical loudness to an average listener. Obviously, the level for the 50 Hz tone is 30 dB higher than the level of the 60 Hz tone, which means that we are much less sensitive to the 50 Hz tone. Based on the equal loudness contours, *loudness level*, in *phons*, is introduced. It is always referenced to a pure tone at 1 kHz. The loudness level of a pure tone (at any frequency) is defined as the level of a 1 kHz tone that has identical loudness to the given tone for an average listener. For the above example, the loudness of the 50 Hz pure tone is 30 phons, which means it is as loud as a 30 dB pure tone at 1 kHz. The lowest curve marked with "minimum audible" is the hearing threshold. Although many normal listeners can hear tones weaker than this threshold at some frequencies, on average, it is a good estimate of a minimum audible limit. The tones louder than the curve of 120 phons will cause pain and hearing damage.

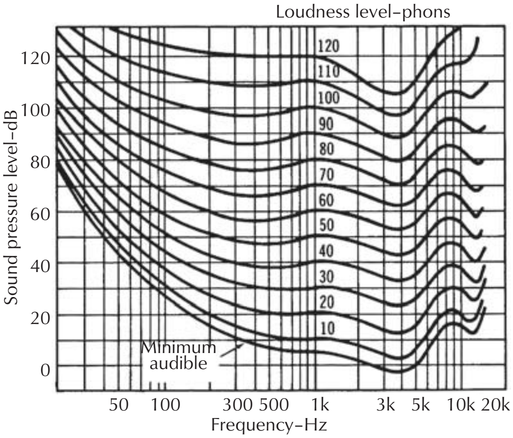


Figure 3-13. Equal loudness contours for pure tones in a frontal sound field for humans of average hearing acuity determined by Robinson and Dadson. The loudness levels in phons correspond to the sound pressure levels at 1000 Hz. (ISO Recommendation 226).

The equal loudness contours also show that human hearing is most sensitive around 4 kHz (which is where the hearing damage due to loud soundsfirst happens), less sensitive to high frequencies, and much less sensitive for very low frequencies (which is why a subwoofer has to be very powerful to produce strong bass, the price of which is the masking of mid-and high-frequencies and potential hearing damage). A study of this family of curves tells us why treble and bass frequencies seem to be missing or down in level when favorite recordings are played back at low levels.[26]

One might notice that for high frequencies above 10 kHz, the curves are nonmonotonic for low levels. This is due to the second resonant mode of the ear canal. Moreover, at low frequencies below 100 Hz, the curves are close to each other, and the change of a few dB can give you the feeling of more than 10 dB of dynamic change at 1 kHz. Furthermore, the curves are much flatter at high levels, which unfortunately encouraged many to listen to reproduced music at abnormally high levels, again causing hearing damage. Actually, even if one wanted to have flat or linear hearing, listening at abnormally high levels might not be wise, because the frequency selectivity of our auditory system will be much poorer, leading to much greater interaction between various frequencies. Of course, one limitation of listening at a lower level is that, if some frequency components fall below the hearing threshold, then they are not audible. This problem is especially important for people who have already lost some acuity at a certain frequency, where his or her hearing threshold is much higher than normal. However, in order to avoid further damage of hearing, and in order to avoid unnecessary masking effect, one still might consider listening at moderate levels.

The loudness level considers the frequency response of our auditory system, and therefore is a better scale than the sound pressure level to account for loudness. However, just like the sound pressure level is not a scale for loudness, the loudness level does not directly represent loudness, either. It simply references the sound pressure level of pure tones at other frequencies to that of a 1 kHz pure tone. Moreover, the equal loudness contours were achieved with pure tones only, without consideration of the interaction between frequency components—e.g., the compression within each auditory filter. One should be aware of this limit when dealing with broadband signals, such as music.

3.8.2 Level Measurements with A-, B-, and C-Weightings

Although psychoacoustical experiments give better results on loudness, practically, level measurement is more convenient. Because the equal loudness contours are flatter at high levels, in order to make level measurements somewhat representing our loudness perception, it is necessary to weight frequencies differently for measurements at different levels. Fig. 3-14 shows the three widely used weighting functions.[27] The A-weighting level is similar to our hearing at 40 dB, and is used at low levels; the B-weighting level represents our hearing at about 70 dB; and the C-weighting level is more flat, representing our hearing at 100 dB, and thus is used at high levels. For concerns on hearing loss, the A-weighting level is a good indicator, although hearing loss often happens at high levels.

3.8.3 Loudness in Sones

Our hearing for loudness is definitely a compressed function (less sensitive for higher levels), giving us both sensitivity for weak sounds and large dynamic range for loud sounds. However, unlike the logarithmic scale (dB) that is widely used in sound pressure level, experimental evidence shows that loudness is actually a power law function of intensity and pressure as shown in Eq. 3-3.

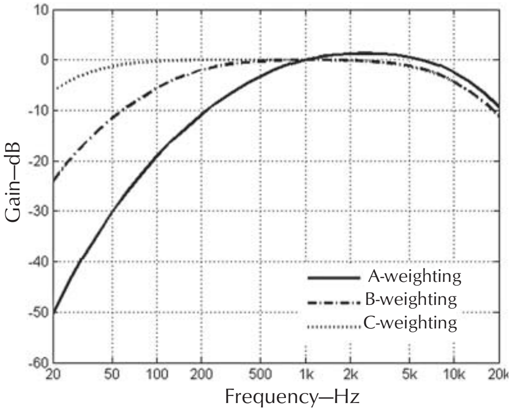


Figure 3-14. Levels with A-, B-, and C-weightings. (Reference 27.)

$$\begin{aligned} Loudness &= k \times I^{\alpha} \\ &= k' \times p^{2\alpha} \end{aligned} \tag{3-3}$$

where,

k and k' are constants accounting for individuality of listeners,

I is the sound intensity,

p is the sound pressure,

α varies with level and frequency.

The unit for loudness is *sones*. By definition, one sone is the loudness of a 1 kHz tone at a loudness level of 40 phons, the only point where phons and SPL meet. If another sound sounds twice as loud as the 1 kHz tone at 40 phons, it is classified as 2 sones, etc. The loudness of pure tones in sones is compared with the *SPL* in dB in Fig. 3-15. The figure shows that above 40 dB, the curve is a straight line, corresponding to an exponent of about 0.3 for sound intensity and an exponent of 0.6 for sound pressure as in Eq. 3-3. The exponent is much greater for levels below 40 dB, and for frequencies below 200 Hz (which can be confirmed by the fact that the equal loudness contours are compact for frequencies below 200 Hz on Fig. 3-13).

One should note that Eq. 3-3 holds for not only pure tones, but also bandpass signals within an auditory filter (critical band). The exponent of 0.3 (<1) indicates compression within the filter. However, for a broadband signal that is wider than one critical bandwidth, Eq. 3-3 holds for each critical band, and the total loudness is simply the sum of loudness in each band (with no compression across critical bands).

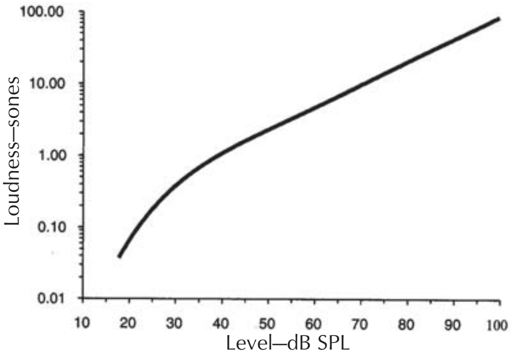


Figure 3-15. Comparison between loudness in sones and loudness level in phons for a 1 kHz tone. (Plack, Reference 15, p118, data from Hellman, Reference 28.)

3.8.4 Loudness versus Bandwidth

Due to less compression across critical bands, broadband sounds, such as a rocket launching or a jet aircraft taking off, seem to be much louder than pure tones or narrow bands of noise of the same sound pressure level. In fact, as in the example in Section 3.3.3, increasing the bandwidth does not increase loudness until the critical bandwidth is exceeded. Beyond that point multiple critical bands are excited, and the loudness increases markedly with increase in bandwidth because of less compression across critical bands. For this reason, the computation of loudness for a wide band sound must be based on spectral distribution of energy. Filters no narrower than critical bands are required and ⅓ octave filters are commonly used.

3.8.5 Loudness of Impulses

Life is filled with impulse-type sounds: snaps, pops, crackles, bangs, bumps, and rattles. For impulses or tone bursts with duration greater than 100 ms, loudness is independent of pulse width. The effect on loudness for pulses shorter than 200 ms is shown in Fig. 3-16. This curve shows how much higher the level of short pulses of noise and pure tones must be to sound as loud as continuous noise or pure tones. Pulses longer than 200 ms are perceived to be as loud as continuous noise or tones of the same level. For the shorter pulses, the pulse level must be increased to maintain the same loudness as for the longer pulses. Noise and tonal pulses are similar in the level of increase required to maintain the same loudness. Fig. 3-16 indicates that the ear has a time constant of about 200 ms, confirming the time window on the order of 100 ms, as discussed in Section 3.7. This means that band levels should be measured with RMS detectors having integration times of about 200 ms. This corresponds to the FAST setting on a sound level meter while the SLOW setting corresponds to an integration time of 500 ms.

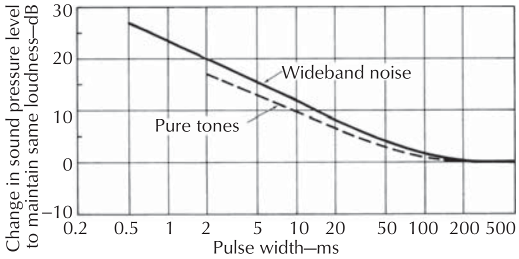


Figure 3-16. Short pulses of sound must be increased in level to sound as loud as longer pulses.

3.8.6 Perception of Dynamic Changes

How sensitive is our hearing of dynamic changes? In other words, how much intensity or level change will lead to a perception of change of loudness? To discuss this kind of problem, we need the concept of just-noticeable difference (JND), which is defined as the minimum change that can be detected. Weber's Law states that the JND in intensity, in general and not necessarily for hearing, is proportional to the overall intensity. If Weber's Law holds, the Weber fraction in dB as defined in Eq. 3-4 would be a constant, independent of the overall intensity and the overall level.

$$\begin{aligned} \textit{Weber fraction} \text{ in dB} &= 10\log\left(\frac{\Delta I}{I}\right) \qquad (3\text{-}4) \\ &= \textit{constant} \end{aligned}$$

where,

I is the intensity,

ΔI is the JND of intensity.

Please note that the Weber fraction in dB is *not* the JND of SPL (ΔL), which can be calculated according to Eq. 3-5.

$$\Delta L = 10\log\left(1 + \frac{\Delta I}{I}\right) \qquad (3\text{-}5)$$

If ΔI is much smaller than I, Eq. 3-5 is approximately

$$\Delta L = 4.35\left(1 + \frac{\Delta I}{I}\right) \qquad (3\text{-}6)$$

Fig. 3-17 shows the measurement of the Weber fraction for broadband signals up to 110 dB SPL.[29] Above 30 dB, the Weber fraction in dB is indeed a constant of about −10 dB, corresponding to a JND (ΔL) of 0.4 dB. However, for weak sounds below 30 dB, the Weber fraction in dB is higher, and can be as high as 0 dB, corresponding to a JND (ΔL) of 3 dB. In other words, our hearing is less sensitive (in level) for dynamic changes of sounds weaker than 30 dB. Interestingly, when measuring with pure tones, it was found that the Weber fraction is slightly different from the broadband signals.[30] This phenomenon is known as the *near-miss Weber's Law*. Fig. 3-17 includes a more recent measurement for pure tones,[31] which demonstrates that the Weber fraction gradually decreases up to 85 dB SPL and can be lower than −12 dB, corresponding to a JND (ΔL) less than 0.3 dB. The near-miss Weber's Law for pure tones is believed to be associated with the broad excitation patterns across frequency at high levels.[32]

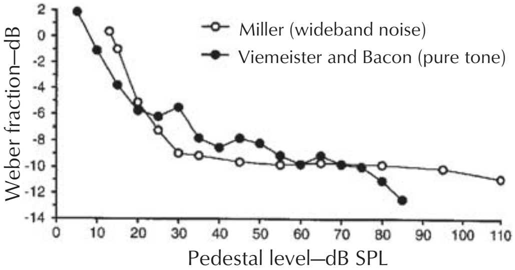


Figure 3-17. Just-noticeable difference (JND) for a broadband noise and for a 1 kHz tone. (After Plack, Reference 15; data from Miller, Reference 29, and Viemeister and Bacon, Reference 31.)

3.9 Pitch

Pitch seems to be a very clear concept, and yet it is very hard to give an accurate definition. The definition by the American National Standards Institute (ANSI) is as follows: "Pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high."[33] Like loudness, pitch is a subjective quantity. The ANSI standard also states: "Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus."[33]

Roughly speaking, the sounds we perceive as pitch are musical tones produced by a musical instrument (except for percussion instruments) and human voice. It is either a pure tone at a certain frequency, or a complex tone with certain fundamentals and a series of harmonics whose frequencies are multiples of the fundamental frequency. For example, when a violin is playing a tone of concert A (440 Hz), the spectrum includes not only the frequency of 440 Hz but also the frequencies of 880 (2 × 440) Hz, 1320 (3 × 440) Hz, and 1760 (4 × 440) Hz, etc.

To perceive a pitch, the sound must be able to match with a pure tone, i.e., a listener must be able to adjust the frequency of a pure tone to produce an identical pitch as the given sound. An opposite example is as follows. When one hits a small drum, it might sound higher than a bigger drum. However, normally one cannot match the sound that a drum produces with a pure tone. The exception, of course, would be a tympani or a steel drum. Therefore, the sound that most drums make does not result in the perception of pitch. Another attribute of pitch is that, if a sound has pitch, one can use it to make a melody. One could use a frequency generator to produce a pure tone at a frequency of 10 kHz, and one could match it with another tone by listening to the beats. However, it would not be perceived as a tone, and it could not be used as part of a melody; therefore it would not be thought of as having pitch.[34] This will be discussed further in Section 3.9.3.

3.9.1 The Unit of Pitch

The unit of *mel* is proposed as a measure of the subjective quantity of pitch.[35] It is always referenced to a pure tone at 1 kHz at 40 dB above a listener's threshold, which is defined as 1000 mels. If another sound produces a pitch that sounds two times as high as this reference, it is considered to be 2000 mels, etc. Fig. 3-18 shows the relationship between pitch in mels and frequency in Hz. The frequency axis in Fig. 3-18 is in logarithmic scale. However, the curve is not a straight line, indicating that our pitch perception is not an ideal logarithmic scale with respect to frequency in Hz. This relationship is probably more important for melodic intervals (when notes are played sequentially) than for chords (when notes are played simultaneously). In a chord, in order to produce a clean harmony, the notes have to coincide with the harmonics of the root note; otherwise, beats will occur, sounding out of tune. In the music and audio industry, it is much more convenient to use frequency in Hz or the unit of *cent* based on the objective quantity of frequency.

Because our hearing is approximately a logarithmic scale on frequency—e.g., doubling frequency transposes a musical note to an octave higher—musical intervals between two tones can be described objectively in the unit of cent as defined by Eq. 3-7.

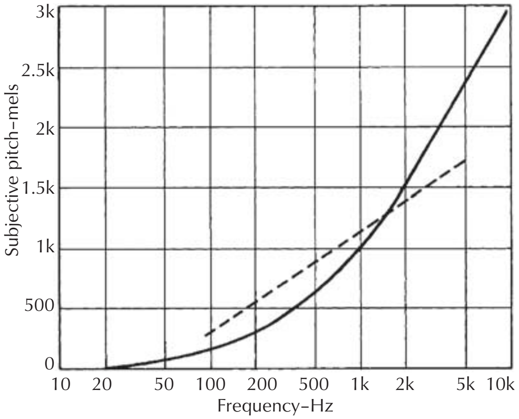


Figure 3-18. The relationship between frequency, a purely physical parameter, and pitch, a subjective reaction to the physical stimulus. (After Stevens et al., Reference 35.)

$$\begin{aligned} \textit{Musical Interval} \text{ in cents} &= 1200\log_2\left(\frac{f_2}{f_1}\right) \\ &= \frac{1200}{\log_{10}2} \times \log_{10}\left(\frac{f_2}{f_1}\right) \end{aligned} \tag{3-7}$$

where,

f_1 and f_2 are the fundamental frequencies of the two tones.

Thus, a semi-tone on a piano (equal temperament) is 100 cents, and an interval of an octave is 1200 cents. Using the unit of cent, one can easily describe the differences among various temperaments (e.g., equal temperament, Pythagorean scale, Just-tuning, etc.).

3.9.2 Perception of Pure and Complex Tones

How does our brain perceive pitch? The basilar membrane in the inner ear functions as a frequency analyzer: pure tones at various frequencies will excite specific locations on the basilar membrane. This would seem to suggest that the location of the maximum excitation on the basilar membrane determines the pitch. Actually, the process is much more complicated: besides the place coding, there is also temporal coding, which accounts for the time interval between two adjacent neural spikes. The temporal coding is necessary for perceiving the pitch of complex tones, the virtual pitch with missing fundamentals, etc.[36,37] The theories based on place coding and temporal coding have been proposed to explain the origin of perception of pure and complex pitches. For either the place theory or the temporal theory, there is experimental evidence supporting and dis-favoring it. As we develop our knowledge, we will probably understand more about when each coding takes place.

When hearing a complex tone, we are tolerant with harmonics slightly mistuned.[38] For instance, if three frequencies of 800, 1000, and 1200 Hz (i.e., the 4th, 5th and 6th harmonics of a fundamental of 200 Hz) are combined and presented to a listener, the pitch perceived is 200 Hz. If all of them are shifted upward by 30 Hz—i.e., 830, 1030, and 1230 Hz—the fundamental, theoretically, is now 10 Hz, and those three components are now the 83rd, 103rd, and 123rd harmonics of the fundamental of 10 Hz. However, when playing this mistuned complex tone, listeners can hear a clear pitch at 206 Hz, which matches the middle frequency—i.e., 1030 Hz—the 4th harmonic of the fundamental. Although the other two frequencies are slightly mistuned (in opposite directions) as harmonics, the pitch is very strong.

It is worth mentioning that pitch recognition is an integrated process between the two ears, in other words, it is a binaural process. When two harmonics of the same fundamental are presented to each ear independently, the listener will hear the pitch at the fundamental frequency, not as two pitches, one in each ear.[39]

3.9.3 Phase-Locking and the Range of Pitch Sensation

What is the range of the fundamental frequency that produces a pitch? Is it the same as the audible range from 20 Hz to 20 kHz? The lowest key in a piano is 27.5 Hz, not too far from the lowest limit. However, for the high limit, it is only about 5 kHz. Because the pitch perception requires temporal coding, the auditory neurons have to fire at a certain phase of each cycle, which is called *phase-locking*. Unfortunately, the auditory system is not able to phase-lock to frequencies above 5 kHz.[40] This is why the highest note on a piccolo, which is the highest pitch in an orchestra, is 4.5 kHz, slightly lower than 5 kHz. Notes with fundamentals higher than 5 kHz are not perceived as having pitch and cannot be used for musical melodies. One can easily confirm this statement by transposing a familiar melody by octaves: when the fundamental is above 5 kHz, although one can hear something changing, the melody cannot be recognized any more.

3.9.4 Frequency Difference Limen

The frequency difference limen is another way of saying "the just-noticeable difference in frequency." It is the smallest frequency difference that a listener can discriminate. Experiments with pure tones of duration of 500 ms show that, for levels higher than 10 dB above threshold, between 200 Hz and 5 kHz, the frequency difference limen is less than 0.5% of the given frequency, corresponding to 9 cents[41] (about 1⁄10 of a semi-tone).

3.9.5 Dependence of Pitch on Level

Pitch can be affected by level, however, the influence is not universal across frequency. At frequencies below 1 kHz, the pitch decreases as level increases; whereas at frequencies above 3 kHz, the pitch increases with increasing level; and at frequencies between 1 and 3 kHz, varying the level has little effect on pitch. This is known as *Stevens rule*. Terhardt et al. summarized several studies of level dependence of pure tones and came up with the following equation for an average listener.[42]

$$100 \times \frac{p-f}{f} = 0.02(L-60)\left(\frac{f}{1000} - 2\right) \tag{3-8}$$

where,

f is the frequency in Hz of a pure tone at a sound pressure level of L in dB,

p is the frequency of a pure tone at 60 dB SPL that matches the pitch of the given tone f.

3.9.6 Perfect Pitch

Some people, especially some musicians, develop perfect pitch, also known as *absolute pitch*. They can identify the pitch of a musical tone without help from an external reference like a tuner—i.e., they have established an absolute scale of pitch in their heads. Some of them describe the feeling of certain note analogous to a certain color. Some believe that one can establish a sensation of perfect pitch if he or she has a lot of experience listening to music on certain keys (which is normally due to musical training) before the age of 4. It is fair to state that having perfect pitch is not a requirement for a fine musician. Other than the advantage of tuning musical instruments or singing without a tuning device, there is no evidence that a person with perfect pitch would sing more accurately in tune. With the help of a tuner or accompaniment, a good musician without perfect pitch would do just as well. There is, however, a disadvantage due to age effect. For senior persons (especially those above the age of 65), the pitch scales are often shifted so that they would hear music that is being played normally to sound out of tune. This might not be noticeable for a person without perfect pitch. However, for a senior musician with perfect pitch, he or she might find it to be annoying when perceiving everyone else in the orchestra playing out of tune. He or she has to live with it, because he or she might be the only one in the orchestra playing out of tune.

3.9.7 Other Pitch Effects

Pitch is mostly dependent on the fundamental frequency, and normally the fundamental frequency is one of the strongest harmonics. However, the fundamental of a complex tone can be missing or masked with a narrowband noise, while still producing a clear pitch.[43] The pitch produced without fundamental is called *virtual pitch*, and it is evidence favoring the temporal theory over the place theory. The waveform of a virtual pitch bears identical period as a normal complex tone including the fundamental at the same frequency.

When listening to a broadband signal with a certain interaural phase relationship, although listening with one ear does not produce a pitch, when listening with both ears, one can hear a pitch on top of the background noise. These kind of pitches are called "binaural pitches."[44,45]

3.10 Timbre

Timbre is our perception of sound color. It is that subjective dimension that allows us to distinguish between the sound of a violin and a piano playing the same note. The definition by the American Standards Association states that the timbre is "that attribute of sensation in terms of which a listener can judge that two sounds having the same loudness and pitch are dissimilar," and "timbre depends primarily upon the spectrum of the stimulus, but it also depends upon the waveform, the sound pressure, the frequency location of the spectrum, and the temporal characteristics of the stimulus."[17] Each sound has its unique spectrum. For musical instruments, the spectrum might be quite different for different notes, although they all sound like tones produced by the same instrument. The timbre of a sound produced in a concert hall may even vary with listener position because of the effects of air absorption and because of the frequency-dependent absorption characteristics of room surfaces.

It is worth noting that, in order to more completely describe timbre, both amplitude and phase spectra are necessary. As the example in Section 3.5 shows, although a white noise and an impulse have identical amplitude spectra, they sound quite differently due to the difference in the phase spectra. Sometimes the onset and offset of a tone might be important for timbre (e.g., the decay of a piano tone). Thus, along with the consideration of the time window of human hearing (on the order of 100 ms), the most accurate description of a timbre would be a spectrogram (i.e., the spectrum developing with time), as shown in Fig. 3-19.

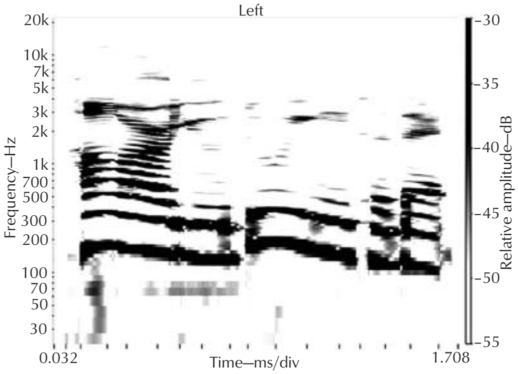


Figure 3-19. An example of a spectrogram: male voice of "How are you doing today?" The vertical axis is frequency, the horizontal axis is time, and the darkness of a point represents the level of a particular frequency component at a given time.

3.11 Binaural and Spatial Hearing

What is the advantage of having two ears? One obvious advantage is a backup: if one ear is somehow damaged, there is another one to use, a similar reason to having two kidneys. This explanation is definitely incomplete. In hearing, having two ears gives us many more advantages. Because of having two ears, we can localize sound sources, discriminate sounds originated from different locations, hear conversations much more clearly, and be more immune to background noises.

3.11.1 Localization Cues for Left and Right

When a sound source is on the left with respect to a listener, it is closer to the left ear than to the right ear. Therefore the sound level is greater at the left ear than that at the right ear, leading to an *interaural level difference* (ILD). Sometimes, people also use the *interaural intensity difference* (IID) to describe the same quantity. Moreover, because the sound wave reaches the left ear earlier than the right ear, there is an *interaural time difference* (ITD) between the two ears. However, the auditory neurons do not directly compare ITD, and instead, they compare the *interaural phase difference* (IPD). For a pure tone, ITD and IPD are linearly related. The ITD and IPD are also referred to as the *interaural temporal difference*. In summary, for localization of left and right, there are two cues—i.e., the ILD and ITD cues.

Adjusting either the ILD or the ITD cues can affect sound localization of left and right. In reality, both of those cues vary. There are limits for both cues. In order to better localize using the ILD cues, the interaural differences should be greater. Because of diffraction around the head, at frequencies below 1 kHz, the levels at both ears are similar, leading to small ILD cues. Therefore ILD cues are utilized at high frequencies, when the head shadow has a big effect blocking the contralateral ear (the one not pointing at the source). On the other hand, there is a limit for the *ITD* cues as well. At frequencies above 700 Hz, the *IPD* of a source at extreme left or right would exceed 180°. For a pure tone, this would lead to confusion: a tone far to the right might sound to the left, Fig. 3-20. Since we care most about the sound sources in front of us, this limit of 700 Hz can be extended upward a little bit. Furthermore, with complex signal with broad bandwidth, we can also use the time delay (or phase difference) of the low-frequency modulation. In general, the frequency of 1.2 kHz, (or a frequency range between 1 and 1.5 kHz) is a good estimate for a boundary, below which ITD cues are important, and above which ILD cues are dominant.

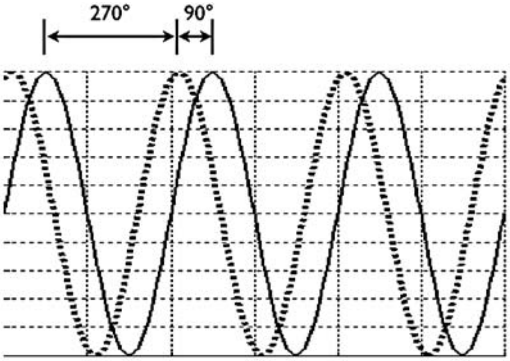


Figure 3-20. Confusion of interaural phase difference (IPD) cues at high frequencies. The dashed curve for the left ear is lagging the solid curve for the right ear by 270°, but it is confused as the left ear is leading the right ear by 90°.

In recording, adjusting the *ILD* cues is easily achieved by panning between the left and right channels. Although adjusting ITD cues also move the sound image through headphones, when listening through loudspeakers, the ILD cues are more reliable than ITD cues with respect to the loudspeaker positions.

3.11.2 Localization on Sagittal Planes

Consider two sound sources, one directly in front of and one directly behind the head. Due to symmetry, the ILD and ITD are both zero for those sources. Thus, it would seem that, by using only ILD and ITD cues, a listener would not be able to discriminate front and back sources. If we consider the head to be a sphere with two holes at the ear-positions (the spherical head model), the sources producing a given ITD all locate on the surface of a cone as shown in Fig. 3-21. This cone is called the *cone of confusion*.[46] If only ITD cues are available for a listener with a spherical head, he or she would not be able to discriminate sound sources on a cone of confusion. Of course, the shape of a real head with pinnae is different from the spherical head, which changes the shape of the cone of confusion, but the general conclusion still holds. When the ILD cues are also available, due to diffraction of the head (i.e., the head shadow), the listener can further limit the confusion into a certain cross-section of the cone (i.e., the dark “donut” in Fig. 3-21). That is the best one can do with ILD and ITD cues. However, in reality, most people can easily localize sound sources in front, in the back, and above the head, etc, even with eyes closed. We can localize sources in a *sagittal plane* (a vertical plane separating the body into, not necessarily equal, left and right parts) with contribution of the asymmetrical shape of our pinnae, head, and torso of our upper body. The pinnae are asymmetrical when looked at from any direction. The primary role of the pinna is to filter or create spectral cues that are virtually unique for every angle of incidence. Different locations on the cone of confusion will be filtered differently, producing spectral cues unique to each location.

The common way of describing the spectral cue for localization is the *head-related transfer function* (HRTF), an example of which is shown in Fig. 3-22. It is the transfer function (gain versus frequency), illustrating the filtering feature of the outer ear, for each location in space (or, more often, for each angle of incidence). Nowadays, with probe microphones inserted close to the eardrum, HRTF can be measured with high accuracy. Once it is obtained for a given listener, when listening to a recording made in an anechoic chamber convolved with the proper HRTF, one can “cheat” the auditory system and make the listener believe that the recording is being played from the location corresponding to the HRTF.

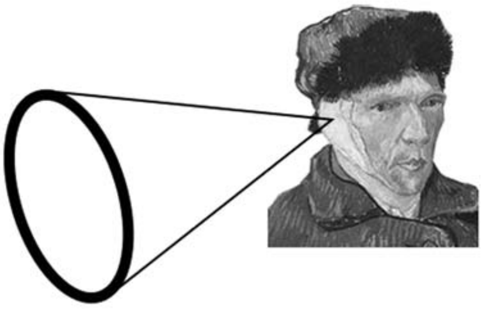


Figure 3-21. Cone of confusion for a spherical head with two holes at the ear positions. If only *ITD* cues are available, the listener cannot discriminate positions on the surface of the cone of confusion, corresponding to a given *ITD*. If *ILD* cues are also available, due to the diffraction of the head, the listener can further limit the confusion range into a circle (the dark "donut" on the figure).

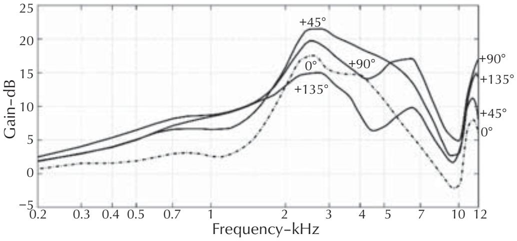


Figure 3-22. Head-related transfer functions. Each curve shows the filtering feature (i.e., the gain added by the external ear at each frequency) of an incident angle. This figure shows the orientations in the horizontal plane. The angles are referenced to the medial sagittal plane, ipsilateral to the ear. The angle of 0° is straight ahead of the subject.

There are two challenges when using spectral cues. The first is discriminating between the filtering feature and the spectrum of the source. For instance, if one hears a notch around 9 kHz, it might be due to an HRTF, or the original source spectrum might have a notch around 9 kHz. Unfortunately there is no simple way to discriminate between them. However, for a familiar sound (voice, instruments, etc.) with a spectrum known to the auditory system, it is easier to figure out the HRTFs and thus easier to localize the source than an unknown sound. If one has trouble discriminating sounds along the cone of confusion, one can use the cues of head motion. For example, suppose a listener turns his or her head to the left. If the source moves to the right, the source is in front; whereas if the source moves farther to the left, it must be in the back. The second challenge is the individuality of HRTFs. No two people share the same pinna and head shape, and we have learned our own pinnae and head size/shape over years of experience. If one listens to sounds convolved with the HRTFs of someone else, although the left–right localization will be good, there will be a lot of front–back confusion,[47] unless the listener's head and ears happen to be similar in size and shape to those whose HRTF is measured.[48] The human binaural system is remarkably adaptive. Experiments with ear molds[49] show that, if a subject listens exclusively through another set of ears, although there is originally a lot of front–back confusion, in about 3 weeks, the subject will learn the new ears and localize almost as well as with their original ears. Instead of forgetting either the new or the old ears, the subject actually memorizes both sets of ears, and becomes in a sense bilingual, and is able to switch between the two sets of ears.

3.11.3 Externalization

Many listeners prefer listening to music through loudspeakers instead of through headphones. One of the reasons is that when listening through headphones, the pinnae are effectively bypassed, and the auditory system is not receiving any of the cues that the pinnae produce. Over headphones, the instruments and singers' voices are all perceived or localized inside the head. When listening through loudspeakers, although the localization cues are not perfect, the sounds are externalized if not localized, somewhat more naturally. If, however, music playing through the headphones includes the HRTFs of the listener, he or she should be able to externalize the sound perfectly.[50] Algorithms are available to simulate 3D sound sources at any location in free field and in a regular room with reverberation. The simulation is accurate to up to 16 kHz, and listeners cannot discriminate between the real source and the virtual (simulated) sound.[51,52] An inconvenience nevertheless is that the system has to be calibrated to each listener and each room. In 1985, Jones et al.[53] devised a test for stereo imagery utilizing a reverberator developed at the Northwestern University Computer Music Studio. The reverberator utilized HRTFs to create very compelling simulations of 3D space and moving sound sources within 3D space. The test by Jones et al.,[53] called LEDR (Listening Environment Diagnostic Recording) NU™, contained sound examples that moved in very specific sound paths. When played over loudspeaker systems that were free from phase or temporal distortions and in environments free from early reflections, the paths were perceived as they were intended. In the presence of early reflection or misaligned crossovers or drivers, the paths are audibly corrupted.

3.11.4 Precedence (Hass) Effect

When two clicks are presented simultaneously to a listener, one on the left and one on the right, the listener would perceive a click in front—i.e., average the localization cues of the two clicks. However, if one of the clicks is delayed (up to 5 ms) compared to the other, the listener still perceives them as one fused click but will localize the fused image with cues of the first click only and ignore the localization cue of the later one. For delays longer than 5 ms, the listener will hear two distinct clicks instead of one fused click. For speech, music or other complex signals, this upper limit can be increased to about 40 ms. This phenomenon that the auditory system localizes on the first arrival is called *precedence effect*, or *Haas effect*.[54,55]

The precedence effect has very practical uses in audio. For example, in a large church, it may not be possible or practical to cover the entire church from one loudspeaker location. One solution is to place a primary loudspeaker in the front of the church, with secondary loudspeakers along the side walls. Because of the precedence effect, if the signal to the loudspeakers along the walls is delayed so that the direct sound from the front arrives at a listener first, the listener will localize the front loudspeaker as the source of the sound, even though most of the content will actually be coming from the loudspeaker to the side, which is much closer to the listener, and may even be operating at a higher level. When such systems are correctly set up, it will sound as though the secondary loudspeakers are not even turned on. Actually turning them off demonstrates exactly how important they are, as without them the sound is unacceptable, and speech may even be unintelligible.

3.11.5 Franssen Effect

The *Franssen Effect*[56] can be a very impressive demonstration in a live room. A pure tone is played through two loudspeakers at two different locations. One loudspeaker plays the tone first and is immediately faded, while the same pure tone is boosted at the other loudspeaker, so that the overall level is not changed significantly, Fig. 3-23. Although the original loudspeaker is not playing at all, most of the audience will still believe that the sound is coming from the first loudspeaker. This effect can last for a couple minutes. One can make this demonstration more effectively by disconnecting the cable to the first loudspeaker, and the audience will still localize the sound to that loudspeaker. The Franssen effect reveals the level of our auditory memory of source locations in a live room.

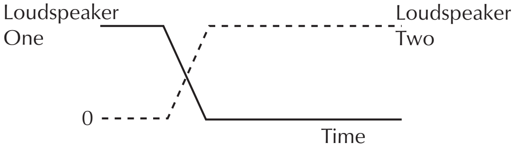


Figure 3-23. Franssen effect (Reference 56). The figure shows the level of two loudspeakers at two difference locations in a live room. Loudspeaker One plays a pure tone first, and is immediately faded. Meanwhile, the same tone played by Loudspeaker Two is boosted, so that the overall level in the room is not changed significantly. After Loudspeaker One stops playing, listeners will still perceive the sound originated from Loudspeaker One, up to a couple minutes.

3.11.6 Cocktail Party Effect and Improvement of Signal Detection

In a noisy environment, such as a cocktail party, many people are talking simultaneously. However, most people have the ability to listen to one conversation at a time, while ignoring other conversations going on around them. One can even do this without turning his or her head to the loudspeaker. As we mentioned earlier, one benefit of binaural hearing is the ability to spatially filter. Because the talkers are spatially separated, our auditory system can filter out unwanted sound spatially. Patients with hearing difficulties usually suffer greatly in a noisy environment because they are unable to pick up an individual's conversation out of the background.

Because the background noise is normally in phase between the two ears, in electronic communication, one can reverse the phase of a signal in one ear and make it out of phase between the ears. Then signal detection is much better due to spatial filtering. So, in general, binaural hearing not only gives us localization ability, but also improves our ability to detect an acoustical signal, especially in a noisy or reverberant environment.

3.11.7 Distance Perception

Distance cues are fairly difficult to replicate. In free field conditions, the sound pressure level will decrease 6 dB with every doubling of the distance between a point source and an observer. Thus reducing the volume should make us feel the source is farther away. In practice, however, we tend to underestimate the distance: the level has to be attenuated by 20 dB in order to give us the perception of a doubled distance.[57] Of course if we do not know how loud the original source is, we do not have an absolute scale based on level.

When a sound source is very far, because the air absorbs high-frequencies more than the low-frequencies, the perceived sound would contain more low-frequency energy, with a darker timbre. This is why thunder far away is just rumble whereas thunder nearby has a crack to it. However, this is a very weak effect[58] and therefore is relatively insignificant for events nearby, which is mostly the case in everyday life.

A more compelling cue for replicating and perceiving distance is adjusting the ratio of the direct to reverberant sound. In real spaces, a sound nearby will not only be louder but also will have a relatively high direct-to-reverberant ratio. As the sound moves away, it gets quieter, and the direct-to-reverberant ratio reduces until critical distance is reached. At critical distance the direct and reverberant levels are equal. Moving a sound source beyond critical distance will not result in an increased sense of distance.

Further Reading

Textbooks on psychoacoustics and auditory physiology are available for various audiences. One might find some of the following books to be helpful:

B. C. J. Moore, “An introduction to the psychology of hearing,” 5th Ed., Academic Press, London (2003).

C. J. Plack, “The sense of hearing,” Lawrence Erlbaum Associates, NJ (2005).

J. O. Pickles, “An introduction to the physiology of hearing,” 2nd Ed., Academic Press, London (1988).

J. D. Durrant and J. H. Lovrinic, “Bases of hearing science,” 3rd Ed., Williams and Wilkins, Baltimore, MD (1995).

W. M. Hartmann, “Signals, sound and sensation,” 5th Ed., Springer, NY (2004).

J. Blauert, “Spatial hearing: The psychophysics of human sound localization,” MIT Press, Cambridge (1997).

H. Fastl and E. Zwicker, “Psychoacoustics: facts and models,” 3rd Ed., Springer, NY (2006).

W. Yost, “Fundamentals of hearing: An introduction,” 5th Ed., Academic Press, NY (2006).

References

1. Douglas Jones, personal communications (2007).
2. J. C. Middlebrooks, “Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency,” *J. Acoust. Soc. Am.*, **106**, 1493-1510 (1999).
3. H. Haas, “Über den Einfluss eines Einfachechos an die Hörsamkeit von Sprache, *Acustica*,” **1**, 49-58 (1951).
4. F. M. Weiner and D. A. Ross, “Pressure distribution in the auditory canal in a progressive sound field,” *J. Acoust. Soc. Am.*, **18**, 401-408 (1946).
5. E. A. G. Shaw, “The external ear,” *Handbook of sensory physiology*, ed. W. D. Keidel and W. D. Neff, vol. **5/1**, pp 455-490, Springer, Berlin (1974).
6. J. D. Durrant and J. H. Loverinic, “Bases of hearing science,” 3rd Ed., p164, Williams and Wilkins, Baltimore, MD (1995).
7. V. Nedzelnitsky, “Sound pressures in the basal turn of the cat cochlea,” *J. Acoust. Soc. Am.*, **68**, 1676-1689 (1980).
8. J. O. Pickles, “An introduction to the physiology of hearing,” 2nd Ed., pp16-20, Academic Press, London (1988).
9. M. Portmann and G. Lacher, “Present management of the reconstruction of the middle ear,” Clinical Otolaryngology, **12** (5), 389–395 (1987).
10. US Patent, “Communication device using bone conduction,” US Patent 6885753.
11. P. W. Carmel and A. Starr, “Acoustic and non-acoustic factors modifying middle-ear muscle activity in waking cats,” *J. Neurophysiol*, **26**, 598-616 (1963).
12. J. D. Durrant and J. H. Loverinic, “Bases of hearing science,” 3rd Ed., p172, Williams and Wilkins, Baltimore, MD (1995).
13. G. von Békésy, “Experiments in hearing,” McGraw-Hill, NY (1960).

14. M. C. Liberman and L. W. Dodds, “Single neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves,” *Hear. Res.*, **16**, 55-74 (1984).
15. C. J. Plack, “The sense of hearing,” Lawrence Erlbaum, NJ (2005).
16. M. A. Ruggero, N. C. Rich, A. Recio, S. S. Narayan, and L. Robles, “Basilar-membrane responses to tones at the base of the chinchilla cochlea,” *J. Acoust. Soc. Am.*, **101**, 2151-2163 (1997).
17. American Standards Association, “Acoustical Terminology SI,” pp1-1960, American Standards Association, NY (1960).
18. H. Fletcher, “Auditory patterns,” *Rev. Mod. Phys.*, **12**, 47-65 (1940).
19. W. Jesteadt, S. P. Bacon, and J. R. Lehman, “Forward masking as a function of frequency, masker level, and signal delay,” *J. Acoust. Soc. Am.*, **71**, 950-962 (1982).
20. A. J. Oxenham and B. C. J. Moore, “Modeling the additivity of nonsimultaneous masking,” *Hear. Res.*, **80**, 105-118 (1994).
21. B. R. Glasberg and B. C. J. Moore, “Derivation of auditory filter shapes from notched-noise data,” *Hear. Res.*, **47**, 103-138 (1990).
22. Zwicker and Terhardt, “Analytical expressions for critical band rate and critical bandwidth as a function of frequency,” *J. Acoust. Soc. Am.*, **68**, 1523-1525 (1980).
23. D. T. Kemp, “Stimulated acoustic emissions from within the human auditory system,” *J. Acoust. Soc. Am.*, **64**, 1386-1391 (1978).
24. D. H. Johnson, “The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones,” *J. Acoust. Soc. Am.*, **68**, 1115-1122 (1980).
25. C. Formby and T. G. Forrest, “Detection of silent temporal gaps in sinusoidal markers,” *J. Acoust. Soc. Am.*, **89**, 830-837 (1991).
26. F. E. Toole, “Loudness – Applications and Implications to Audio,” *dB The Sound Engineering Magazine*, Part I, vol. 7, no. 5, pp.27-30, May 1973; Part II, vol. 7, no. 6, pp.25-28, June 1973.
27. ANSI S1.4-1983, ANSI S1.42-2001, IEC 61672-1:2002.
28. R. P. Hellman, “Growth of loudness at 1000 and 3000 Hz,” *J. Acoust. Soc. Am.*, **60**, 672-679 (1976).
29. G. A. Miller, “Sensitivity to changes in the intensity of white noise and its relation to masking and loudness,” *J. Acoust. Soc. Am.*, **19**, 609-619 (1947).
30. W. J. McGill and J. P. Goldberg, “A study of the near-miss involving Weber’s Law and pure-tone intensity discrimination,” *Percept. Psychophys.*, **4**, 105-109 (1968).
31. N. F. Viemeister and S. P. Bacon, “Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones,” *J. Acoust. Soc. Am.*, **84**, 172-178 (1988).
32. B. C. J. Moore and D. H. Raab, “Pure-tone intensity discrimination: some experiments relating to the ‘near-miss’ to Weber’s Law,” *J. Acoust. Soc. Am.*, **55**, 1049-1054 (1974).
33. American National Standards, “Acoustical Terminology SI,” American National Standards Institute, New York (1994).
34. F. Attneave and R. K. Olson, “Pitch as a medium: a new approach to psychophysical scaling,” *Am. J. Psychol.*, **84**, 147-166 (1971).
35. S. S. Stevens, J. Volkman, and E. B. Newman, “A scale for the measurement of the psychological magnitude of pitch,” *J. Acoust. Soc. Am.*, **8**, 185-190 (1937).
36. J. F. Schouten, “The residue and the mechanism of hearing,” *Proc. Kon. Akad. Wetenschap.*, **43**, 991-999 (1940).
37. J. F. Schouten, “The residue revisited,” *Frequency Analysis and Periodicity Detection in Hearing*, ed. R. Plomp and G. Smoorenburg, A. W. Sijthoff, Leiden, pp41-58 (1970).
38. B. C. J. Moore, B. R. Glasberg, and R. W. Peters, “Relative dominance of individual partials in determining the pitch of complex tones,” *J. Acoust. Soc. Am.*, **77**, 1853-1859 (1985).
39. A. J. M. Houtsma and J. L. Goldstein, “The central origin of the pitch of pure tones: evidence from musical interval recognition,” *J. Acoust. Soc. Am.*, **51**, 520-529 (1972).
40. D. H. Johnson, “The relationship between spike rate and synchrony in responses of auditory nerve fibers to single tones,” *J. Acoust. Soc. Am.*, **68**, 1115-1122 (1980).
41. C. C. Wier, W. Jesteadt, and D. M. Green, “Frequency discrimination as a function of frequency and sensation level,” *J. Acoust. Soc. Am.*, **61**, 178-184 (1977).

42. E. Terhardt, G. Stoll, and M. Seewann, “Algorithm for extraction of pitch and pitch salience from complex tonal signals,” *J. Acoust. Soc. Am.*, **71**, 679-688 (1982).
43. J. C. R. Licklider, “Auditory frequency analysis,” Information Theory, ed. C. Cherry, pp253-268, Academic, NY (1956).
44. E. M. Cramer and W. H. Huggins, “Creation of pitch through binaural interaction,” *J. Acoust. Soc. Am.*, **30**, 413-417 (1958).
45. F. A. Bilsen and J. Raatgever, “On the dichotic pitch of simultaneously presented interaurally delayed white noises. Implications for binaural theory,” *J. Acoust. Soc. Am.*, **108**, 272-284 (2000).
46. A. W. Mills, “Auditory localization,” *Foundations of Modern Auditory Theory*, *Vol. II*, ed. J. V. Tobias, Academic Press, NY (1972).
47. F. L. Wightman and D. J. Kistler, “Headphone simulation of free-field listening. I: stimulus synthesis,” *J. Acoust. Soc. Am.*, **85**, 858-867 (1989).
48. J. C. Middlebrooks, “Individual differences in external-ear transfer functions reduced by scaling in frequency,” *J. Acoust. Soc. Am.*, **106**, 1480-1492 (1999).
49. M. B. Gardner and R. S. Gardner, “Problem of localization in the median plane: effect of pinnae cavity occlusion,” *J. Acoust. Soc. Am.*, **53**, 400-408 (1973).
50. W. M. Hartmann and A. Wittenberg, “On the externalization of sound images,”*J. Acoust. Soc. Am.*, **99**, 3678-3688 (1996).
51. P. X. Zhang and W. M. Hartmann, “Transaural virtual reality applied to front-back localization,” *J. Acoust. Soc. Am.*, **117**, 2485-2486 (2005).
52. P. X. Zhang, “Detection and localization of sounds: virtual tones and virtual reality,” Thesis, Michigan State University, East Lansing (2006).
53. D. R. Jones, W. L. Martens, and G. S. Kendall, “Optimizing control rooms for stereo imagery,” *J. Acoust. Soc. Am.*, **78**, S9 (1985).
54. H. Haas, “The Influence of a Single Echo on the Audibility of Speech,” *Acustica*, vol. 1, no. 2, 1951, in German. (An English translation by Dr. Ingr.K. P. R. Ehrenberg appears in the *Journal of the Audio Engineering Society*, vol. 20, no. 2, pp.146 – 159, March 1972.)
55. R. Y. Litovsky, H. S. Colburn, W. A. Yost, and S. J. Guzman, “The precedence effect,” *J. Acoust. Soc. Am.*, **106**, 1633-1654 (1999).
56. W. M. Hartmann and B. Rakerd, “Localization of sound in rooms IV: The Franssen effect,” *J. Acoust. Soc. Am.*, **86**, 1366-1373 (1999).
57. J. Blauert, “Spatial hearing: The psychophysics of human sound localization,” pp122-123, MIT Press, Cambridge (1997).
58. U. Ingard, “A review of the influence of meteorological conditions on sound propagation,” *J. Acoust. Soc. Am.*, **25**, 405-411 (1953).

This page intentionally left blank

Chapter 4

Acoustical Treatment for Indoor Areas

by Doug Jones

4.1 Noise Criteria 67
4.2 Site Selection 69
  4.2.1 Site Noise Survey 69
  4.2.2 Transmission Loss 71
  4.2.3 Sound Barriers 71
  4.2.4 Sound Transmission Class (STC) 72
4.3 Isolation Systems 74
  4.3.1 Wall Construction 74
  4.3.2 High-Loss Frame Walls 74
  4.3.3 Concrete Block Walls 77
  4.3.4 Concrete Walls 77
  4.3.5 Wall Caulking 77
  4.3.6 Floor and Ceiling Construction 78
  4.3.7 Floor Construction 79
  4.3.8 Summary of Floating Floor Systems 82
  4.3.9 Acoustical Doors 83
  4.3.10 Windows 84
  4.3.11 Transmission Loss of a Compound Barrier 87
  4.3.12 Isolation Systems Summary 89
4.4 Heating, Ventilating, and Air Conditioning (HVAC) Systems for Low Noise 89
  4.4.1 Location of HVAC Equipment 90
  4.4.2 Identification of HVAC Noise Producers 90
  4.4.3 Vibration Isolation 90
  4.4.4 Attenuation of Noise in Ducts 91
  4.4.5 Tuned Stub Noise Attenuators 92
  4.4.6 Plenum Noise Attenuators 92
  4.4.7 Proprietary Silencers 93
  4.4.8 HVAC Systems Conclusion 93
References 93

This page intentionally left blank

The field of room acoustics divides easily into two broad categories; noise control and subjective acoustics. These two branches of acoustics actually have rather little in common. Noise control, to a great extent, must be designed into a project. It is very difficult, if not impossible to retroactively improve the isolation of a room. However, it is often possible to change the way a room sounds subjectively by simply modifying the wall treatment. It is important to keep in mind that noise is a subjective categorization. Sound pressure, sound intensity, and sound transmission can all be measured. Noise is unwanted sound. It is much more difficult to measure and quantify the extent to which any given sound annoys any given individual. To the Harley rider driving his Fat Boy™ past your studio, the exhaust is music to his ears but noise to you. Your music is noise to the therapist trying to conduct a session of a very different sort next door. The music in studio A is noise to the band trying to record in studio B down the hall.

Throughout this chapter, the term *sound room* will be used to designate any room that requires some measure of quiet in order for the room to serve its purpose.

4.1 Noise Criteria

When specifying permissible noise levels, it is customary to use some form of the noise criteria (NC). The beauty of the NC contours is that a spectrum specification is inherent in a single NC number. The NC contours of Fig. 4-1 are helpful in setting a background noise goal for a sound room.[1] Other families of NC contours have been suggested such as the PNC,[2] Fig. 4-2 which adds an additional octave to the low end of the scale, and NR (noise rating), Fig. 4-3, used in Europe. In 1989 Beranek proposed the NCB or Balanced Noise Criteria.[3] The NCB adds the 16 Hz octave band and the slopes of the curves are somewhat modified relative to the NC or PNC curves, Fig. 4-4. Beranek also proposed NCB limits for various applications as shown in Table 4-2.

Considering the spectrum of noise is far superior to using a single, wideband noise level. However, if desired, each NC contour can be expressed as an overall decibel level by adding the sound power in each octave band as in Table 4-1. These overall levels are convenient for rough appraisal of noise levels from a single sound level meter (SLM) reading. For example, if the SLM reads 29 dB on the A-weighting scale for the background noise of a studio, it could be estimated that the NC of that room is close to NC-15 on the assumption that the noise spectrum of that room matched the corresponding NC contour, and that there are no dominant pure tone components.

Table 4-1. Noise Criteria (NC) Overall Levels*

NC Contour	Equivalent Wideband Level (A-weighted)
15	28
20	33
25	38
30	42
35	46
40	50
45	55
50	60
55	65
60	70
65	75

*Source: Rettinger[5]

It is helpful to see recommended NC ranges for recording studios and other rooms compared to criteria applicable to spaces used for other purposes, Table 4-2. The NC goals for concert halls and halls for opera,

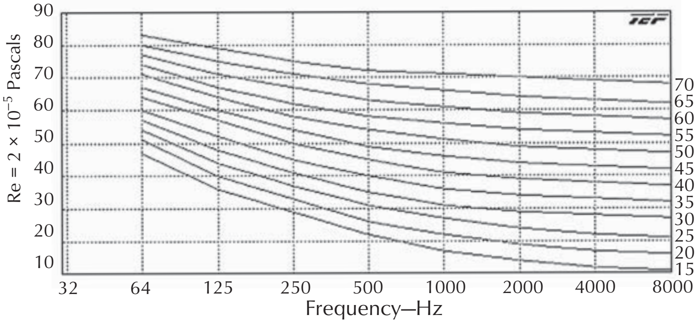


Figure 4-1. Noise criteria (NC) curves. As displayed on Gold Line TEF 20.

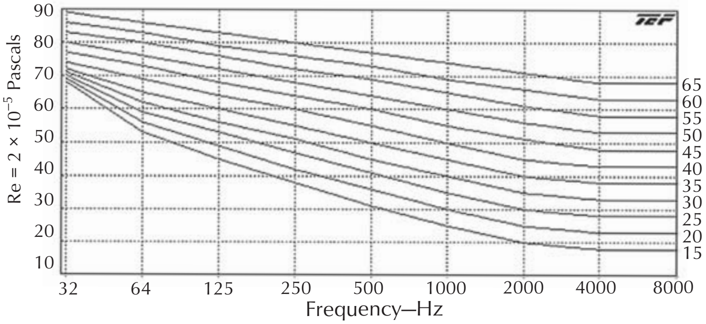


Figure 4-2. PNC. From Gold Line TEF20. Note extra octave.

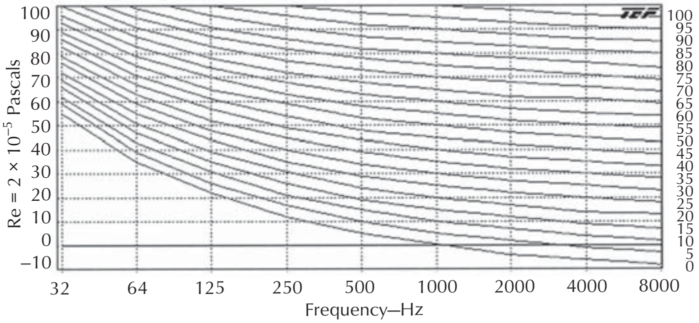


Figure 4-3. NR—The European noise rating curves. From Gold Line TEF 20. Note the extended range.

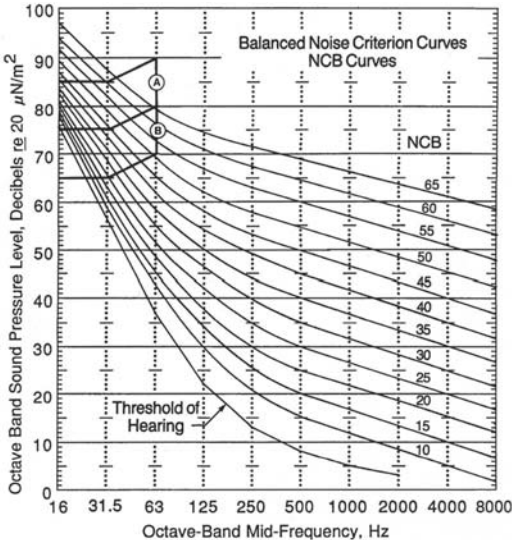


Figure 4-4. Balanced noise-criterion (NCB) curves for occupied room.

musicals, and plays are low to assure maximum dynamic range for music and greatest intelligibility for speech. This same reasoning applies to high-quality listening rooms such as control rooms. For recording studios, stringent NC goals are required to minimize noise pickup by the microphone. Levels below NC-30 are generally considered "quiet," but there are different degrees of quietness. An NC-25 is at the low end of what is expected of an urban residence. This means that

Table 4-2. Noise Criteria Ranges, NC and NCB*

Use of Space	Noise Criteria Range	NCB
Private urban residence, corridors	25 – 35	25 – 40
Private rural residence	20 – 30	na
Hotel rooms	30 – 40	25 – 40
Hospital, private rooms	25 – 35	25 – 40
Hospital, lobby, corridors	35 – 45	40 – 50
Office, executive	30 – 40	30 – 40
Offices, open	35 – 45	35 – 45
Restaurant	35 – 45	35 – 45
Church sanctuary	20 – 30	20 – 30
Concert, opera halls	15 – 25	10 – 15
Studios, recording and sound reproduction	15 – 25	10

*Selected from references 2, 3, and 5

if an NC-25 were met in an urban residence, it is likely that the occupants would perceive it as being quiet. That same NC-25 represents the *upper* limit of what is acceptable for a recording studio. Most recording studios, especially those that record material with wide dynamic content, will require NC-20 or even 15. Levels lower than NC-15 require expensive construction and are difficult to achieve in urban settings.

4.2 Site Selection

Part of meeting a noise goal is the careful selection of a building site, a site that is appropriate for the application and where the NC is achievable and affordable. It is one thing to build a room meeting an NC-15 in a cornfield in central Iowa. It is another thing altogether to build an NC-15 room in downtown Manhattan. When surveying a site, watch for busy roads, especially freeways; elevated, ground level, or underground railroads, busy intersections, airports, and fire stations. When economic or other factors make such a location imperative, allowance must be made for the extra cost of the structure to provide the requisite protection from such noise. When considering space in an existing building, inspect all neighboring spaces and be wary of adjacent spaces that are vacant unless the owner of the sound-sensitive space also controls the vacant space.

Remember that buildings can be very noisy spaces. Sources of noise include elevator doors and motors, heating, ventilating, and air-conditioning equipment; heel taps on hard floors, plumbing, and business machines.

If selecting a plot of land, a limited amount of protection can be achieved by erecting earthen embankments or masonry walls between the structure and the noise source. These are reasonably effective at high frequencies, but low-frequency components of noise whose wavelengths are large relative to the size of the embankment tend to diffract over the top. A stand of dense shrubbery might yield as much as 10 dB of overall attenuation. Physical separation of the proposed structure from the noise source is helpful but limited by the inverse-square law. The 6 dB per distance double rule applies only to point sources in free-field conditions but it is useful for rough estimation. Going from 50 ft to 100 ft (a change of 50 ft) from the source yields the same reduction of noise level as going from 100 ft to 200 ft (a change of 100 ft). Clearly, increasing separation counts most when close to the source. At any given location sites, locating the sound-sensitive rooms on the face of the building away from a troublesome noise source is favorable, especially if no reflective structures are there to reduce the barrier effect.

There are two ways that noise travels from the source to the observer. It is either transmitted through the air—airborne noise—or carried through the structure or the earth—structure borne noise. A highway carrying heavy truck traffic or an overhead or subway railroad, may literally shake the earth to such an extent that large amplitude, low-frequency vibrations of the ground may be conducted to the foundation of the structure and carried to all points within that structure. Even if such vibrations are subsonic, they have been known to shake microphones with good low-frequency response so as to overload low level electronic circuits. Vibration, both subsonic and sonic, is carried with amazing efficiency throughout a reinforced concrete structure. The speed of sound in air is 344 m/s whereas the speed of sound in reinforced concrete, for example, is on the order of 3700 m/s.[5] A large-area masonry wall within a structure, when vibrated at high amplitude, can radiate significant levels of sound into the air by diaphragmatic action. It is possible by using a combination of vibration-measuring equipment and calculations (outside the scope of this treatment) to estimate the sound pressure level radiated into a room via such a structure-borne path. In most cases noise is transmitted to the observer by both air and structure.

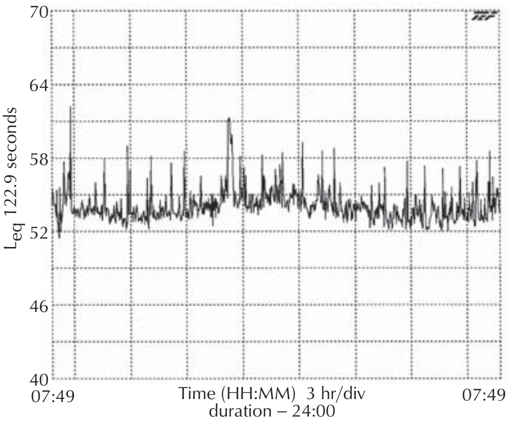
4.2.1 Site Noise Survey

A site survey gives the designer a good idea of the noise levels present at the proposed building site. It is important to know how much noise exists in the immediate environment so that appropriate measures can be taken to reduce it to acceptable levels.

Ambient noise is very complex, a fluctuating mixture of traffic and other noises produced by a variety of human and natural sources. The site noise should be documented with the appropriate test equipment. Subjective approaches are unsatisfactory. Even a modest investment in a studio suite or a listening room justifies the effort and expense of a noise survey of the site which provides the basis for designing walls, floor, and ceiling to achieve the low background noise goals.

One approach to a noise survey of the immediate vicinity of a proposed sound room is to contract with an acoustical consultant to do the work and submit a report. If technically oriented persons are available, they may be able to turn in a credible job if supplied with the right equipment and given some guidance.

The easy way to survey a proposed site is to use one of the more sophisticated microprocessor-based recording noise analyzers available today. There are a number of fine units that are capable of producing reliable and very useful site surveys. Fig. 4-5 is an example of a 24 hour site survey made with the Gold Line TEF 25 running Noise Level Analysis™ software. One can also use a handheld sound level meter (SLM) if outfitted with the appropriate options. Some real-time frequency analyzers such as the Brüel and Kjaer 2143 are also appropriate. It is also acceptable to use a dosimeter such as the Quest Technologies model 300, Fig. 4-6.



L_{min} = 44.4 dB L_{max} = 90.8 dB L_{eq} = 54.3 dB
L_{10} = 57.1 dB L_{50} = 51.5 dB L_{90} = 47.8 dB L_{mean} = 52.1 dB

Channel A Preamp
Gain: 48 dB, Weight: F1t, Output: Off

Figure 4-5. A 24 hour NLA site survey made with the Gold Line TEF 25.

No matter which analyzer is used, the system must be calibrated using a microphone calibrator. The weather conditions, especially temperature and relative humidity, should be noted at the time of calibration. The measuring microphone may be mounted in a weatherproof housing at the desired location with the microphone cable running to the equipment indoors. There are a number of terms which will appear on any display of a noise survey. There will be a series of L_n levels indicated, see Table 4-3. These are called *exceedance* or *percentile levels*. L_{10} refers to the noise level exceeded 10% of the time, L_{50} the level exceeded 50% of the time, l_{90} the level exceeded 90% of the time and so forth. In the United States L_{10} is considered to indicate the *average* maximum level and L_{90} the *average* minimum or background level.[6] Since many noise levels vary dramatically over time, it is useful to have a number which represents the equivalent constant decibel level. The L_{eq}. This is the steady continuous level that would yield the same energy over a given period of time as the measured levels. L_{dn} indicates a 24 hour L_{eq} with 10 dB added to the levels accumulated between 2200 and 0700 h to account for the increased annoyance potential during the nighttime hours. The Community Noise Equivalent Level (CNEL) also is used to document noise levels over a 24 hour period. It differs from the L_{dn} as weighting factors for the evening period between 1900 to 2200 h are included. The L_{eq} for evening hours is increased by 5 dB while the L_{eq} for the nighttime hours is increased by 10 dB.

Quest
Technologies

Figure 4-6. Quest Technologies model 300 dosimeter. Courtesy Quest Technologies.

Table 4-3. Common Level Designations in Noise Surveys

L_{10}	Noise level exceeded 10% of the time
L_{50}	Noise level exceeded 50% of the time
L_{90}	Noise level exceeded 90% of the time
L_{dn}	24 hour L_{eq}
l_{mean}	Arithmetic mean of measured levels

The l_{mean} is the arithmetic mean of the measured levels. L_{min} and L_{max} refer to the lowest and highest measured instantaneous levels, respectively.

Ideally, the site survey should take place over a minimum of 24 hours. A 24 hour observation captures diurnal variations; observations on selected days of the week capture especially noisy events varying from day to day or occurring at certain times during the week.

All of these analyzers, with the exception of the Brüel and Kjaer 2143, will capture only the sound pressure levels over time. The 2143 will also capture the spectrum of the noise as well. If an analyzer such as the 2143 is not available, it is advisable to make a number of measurements of the spectrum as well as the time-stamped level record.

The data collected from the site survey should be combined with projections of the levels anticipated in the sound room and what will be tolerated in adjacent spaces.

4.2.2 Transmission Loss

Once the noise load is known and the desired NC is determined, attention must be given to the design of systems that will provide enough isolation to achieve the goal. Transmission loss is the loss that occurs when a sound goes through a partition or barrier. A higher *TL* number means more loss, i.e., less acoustic energy gets through. If the desired NC or noise limit is known, and the noise load is known, a designer must then design barriers or partitions that have appropriate *TL* to meet the design goal.

$$TL = 10\log\left(\frac{P_{incident}}{P_{transmitted}}\right) \tag{4-1}$$

4.2.3 Sound Barriers

The purpose of a sound barrier is to attenuate sound. To be effective, the barrier must deal with airborne as well as structure-borne noise. Each barrier acts as a diaphragm, vibrating under the influence of the sound impinging upon it. As the barrier vibrates, some of the energy is absorbed, and some is reradiated. The simplest type of barrier is the limp panel or a barrier without any structural stiffness. Approached theoretically, a limp panel should give a transmission loss increase of 6 dB for each doubling of its mass. In the real world, this figure turns out to be nearer 4.4 dB for each doubling of mass. The empirical mass law deduced from real-world measurements can be expressed as

$$TL = 14.5\log M + 23 \tag{4-2}$$

where,

TL is the transmission loss in decibels,

M is the surface density of the barrier in pounds per square foot.

Transmission loss also varies with frequency, even though Eq. 4-1 has no frequency term in it. With a few reasonable assumptions, the following expression can be derived, which does include frequency:[8]

$$TL = 14.5\log(Mf) - 16 \tag{4-3}$$

where,

f is the frequency in hertz.

Fig. 4-7 is plotted from the empirical mass law stated in Eq. 4-3, which is applicable to any surface density and any frequency, as long as the mass law is operating free from other effects.

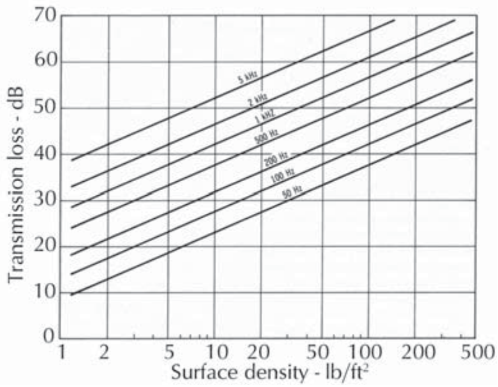


Figure 4-7. The empirical mass law based on real-world measurements of transmission loss. Surface density is the weight of the wall corresponding to a 1 ft^2 wall surface.

From Fig. 4-7 several general conclusions can be drawn. One is that at any particular frequency, the heavier the barrier, the higher the transmission loss. A concrete wall 12 in (30 cm) thick with a surface density of 150 lb/ft^2 (732 kg/m^2) gives a higher transmission loss than a ¼ in (6 mm) glass plate with a surface density of 3 lb/ft^2 (14.6 kg/m^2). Another conclusion is that for a given barrier the higher the frequency, the higher the transmission loss.

The straight lines of Fig. 4-7 give only a partial picture since barrier effects other than limp mass dominate. Fig. 4-8 shows four different regions in the frequency domain of a barrier. At extremely low frequencies, stiffness of the barrier dominates. At somewhat higher frequencies, resonance effects control as the barrier vibrates like a diaphragm. Above a critical frequency, a coincidence effect controls the transmission loss of the barrier. The mass law is an important effect in determining barrier performance, but resonance and coincidence cause significant deviations.

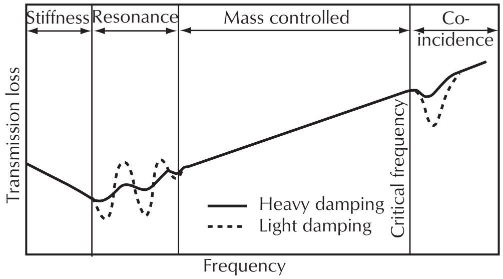


Figure 4-8. The performance of a barrier is divided into four regions controlled by stiffness, resonance, mass, and coincidence.

The low-frequency resonance effect is due to the mechanical resonance of the barrier. For heavier barriers, the resonant-frequency is usually below the audible limit. As the panel vibrates at resonance, there is virtually no transmission loss. At frequencies above resonance, the mass law is in effect, and the function stays fairly linear until the coincidence effect. The coincidence effect occurs when the wavelength of the incident sound coincides with the wavelength of the bending waves in the panel. For a certain frequency and a certain angle of incidence, the bending oscillations of the panel will be amplified, and the sound energy will be transmitted through the panel with reduced attenuation. The incident sound covers a wide range of frequencies and arrives at all angles, but the overall result is that the coincidence effect creates an "acoustical hole" over a narrow range of frequencies giving rise to what is called the c*oincidence dip* in the transmission loss curve. This dip occurs above a critical frequency, which is a complex function of the properties of the material. Table 4-4 lists the critical frequency for some common building materials.

Table 4-4. Critical Frequencies

Material	Thickness (Inches)	Critical Frequency (Hz)
Brick wall	10	67
Brick wall	5	130
Concrete wall	8	100
Glass plate	¼	1600
Plywood	¾	700

*Calculated from Rettinger[8]

4.2.4 Sound Transmission Class (STC)

The noise criterion approach is convenient and valuable because it defines a permissible noise level and spectrum by a single NC number. It is just as convenient and valuable to be able to classify the transmission loss of a barrier versus frequency curve by a single number. The STC or sound transmission class, is a single number method of rating partitions.[9] A typical standard contour is defined by the values in Table 4-5. A plot of the data in Table 4-5 is shown in Fig. 4-9. Only the STC-40 contour is shown in Fig. 4-9, but all other contours have exactly the same shape. It is important to note that the STC is not a field measurement. The field STC, or FSTC, is provided for in ASTM E336-97 annex a1. The FSTC is often 5 dB or so worse than the laboratory STC rating. Therefore a door rated at STC-50 can be expected to perform around STC-45 when installed.

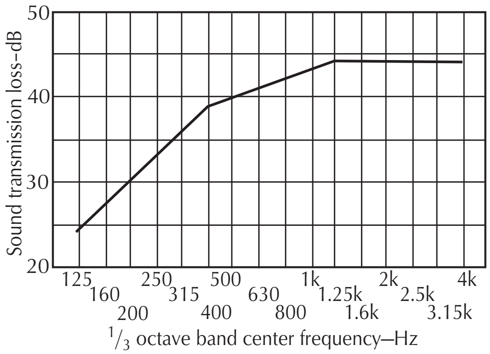


Figure 4-9. The standard shape used in determining the sound transmission class (STC) of a partition (ASTM E413-87).

Nonetheless, the STC provides a standardized way to compare products made by competing manufacturers.

Assume that a TL versus frequency plot of a given partition is at hand and that we want to rate that partition with an STC number. The first step is to prepare a transparent overlay on a piece of tracing paper of the standard STC contour (the STC-40 contour of Table 4-5 and Fig. 4-9) to the same frequency and TL scales as the TL graph. This overlay is then shifted vertically until some of the measured TL values are below the contour and the following conditions are fulfilled:[10]

1. The sum of the deficiencies (i.e., the deviations below the contour) shall not be greater than 32 dB.
2. The maximum deficiency at any single test point shall not exceed 8 dB.

When the contour is adjusted to the highest value that meets these two requirements, the sound transmission class of that partition is the TL value corresponding to the intersection of the contour and the 500 Hz ordinate. An example of the use of STC is given in Fig. 4-10. To determine the STC rating for the measured TL curve shown in Fig. 4-10, the STC overlay is first aligned to 500 Hz and adjusted vertically to read some estimated value, say, STC-44. The difference between the measured TL level and the STC curve is recorded at each of the ⅓ octave points. These data are added together. The total, 47 dB, is more than the 32 dB allowed. The STC overlay is next lowered to an estimated STC-42, and a total of 37 dB results. Lowering the overlay to STC-41 yields a total of 29 dB, which fixes the STC-41 contour as the rating for the TL curve of Fig. 4-10.

Table 4-5. Standard STC Contour

Frequency in Hz	⅓ Octave Sound Transmission Loss in dB	Frequency (Hertz)	⅓ Octave Sound Transmission Loss in dB
125	24	800	42
160	27	1000	43
200	30	1250	44
250	33	1600	44
315	36	2000	44
400	39	2500	44
500	40	3150	44
630	41	4000	44

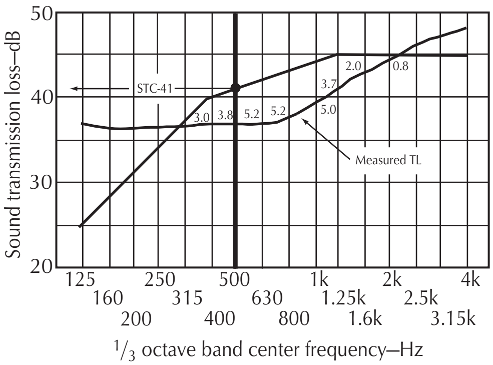


Figure 4-10. The method of determining the single-number STC rating of a barrier from its measured *TL* graph.

The final illustration of STC methods is given in Fig. 4-11. In this case, a pronounced coincidence dip appears at 2500 Hz. This illustrates the second STC requirement, "the maximum deficiency at any single test point shall not exceed 8 dB." This 8 dB requirement fixes the overlay at STC-39, although it might have been considerably higher if only the 32 dB sum requirement applied.

The shape of the standard STC contour may be very different from the measured TL curve. For precise work, using measured, or even expertly estimated, TL curves may be desirable rather than relying on STC single number ratings. Convenience usually dictates use of the STC shorthand system, but it is, at best, a rather crude approximation to the real-world TL curves.

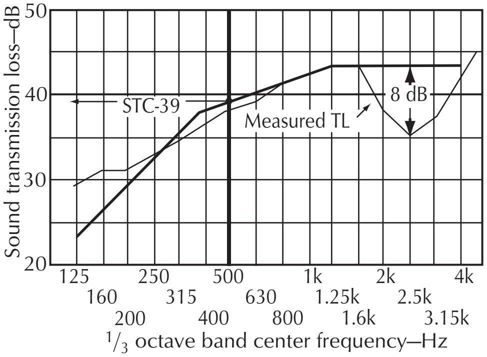


Figure 4-11. The second rule for STC determination that a maximum deficiency of 8 dB is allowed.

Assume that a goal of NC-20 has been chosen for a sound room. The noise survey indicates a noise level and spectrum as shown in Fig. 4-12. What wall construction will bring the noise of Fig. 4-12 down to the NC-20 goal we have set for the interior? Fig. 4-13 shows that a wall having a rating of STC-55 is required. The next step is to explore the multitude of possible wall configurations to meet the STC-55 requirement as well as other needs.

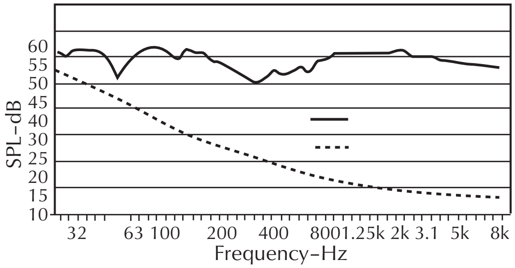


Figure 4-12. Noise spectrum from noise survey.

If the NC curve in Fig. 4-12 is subtracted from the measured noise curve, this will indicate the raw data that indicates the amount of loss needed to achieve the desired NC. This is plotted in Fig. 4-13. The standard STC template is laid over the graph and the needed STC is read opposite the 500 Hz mark.

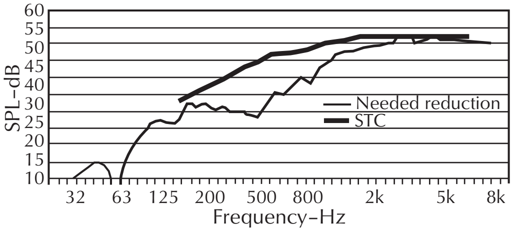


Figure 4-13. The sound barrier attenuation required for the sound room example of Fig. 4-12 is specified here as STC-47.

4.3 Isolation Systems

Isolation systems must be dealt with holistically. One must consider walls, ceilings, floors, windows, doors, etc. as parts of a whole isolation system. Vibration takes every path possible when traveling from one spot to another. For example, if one intends to build a sound room directly below a bedroom of another tenant in a building, one might assume that special attention must be paid to the ceiling. Of course this is correct. However, there are often paths that would permit the vibration to bypass the ceiling. All these flanking paths must be accounted for if isolation between two spaces is desired.

It should be noted that in some parts of the country (most notably California) building codes require seismic engineering. Make sure that the isolation systems that are under consideration do not violate any local seismic codes or require additional seismic restraints. Mason Industries has published a bulletin that is quite instructive in seismic engineering.[11]

4.3.1 Wall Construction

Acoustic partitions are complex entities. As was previously noted, walls exhibit different degrees of isolation in different segments of the spectrum. It is therefore imperative that you know what frequencies you are isolating. (Refer to Fig. 4-8.) The more massive the wall and the more highly damped the material, the fewer the problems introduced by diaphragmatic resonance. In comparing the relative effectiveness of various wall configurations, the mass law offers the most easily accessible rough approximation. However, most practical acoustical partitions actually perform better—that is, they achieve more loss—than what is predicted by the mass law. To assist in the computation of isolation based on mass, the densities of various common building materials are listed in Table 4-6. If an air space is added as in double wall construction, this introduces an element other than mass and generally leads to higher transmission loss.

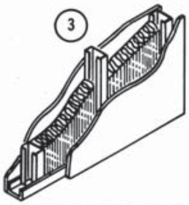
Table 4-6. Building Material Densities

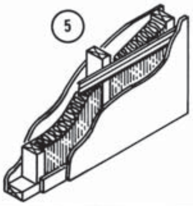
Material (inches)	Density (lb/ft^3)	Surface Density (lb/ft^2)
Brick	120	
4		40.0
8		80.0
Concrete: light wt.	100	
4		33.0
12		100.0
Concrete: dense	150	
4		50.0
12		150.0
Glass	180	
¼		3.8
½		7.5
¾		11.3
Gypsum wallboard	50	
½		2.1
⅝		2.6
Lead	700	
¹⁄₁₆		3.6
Particle Board	48	
¾		1.7
Plywood	36	
¾		2.3
Sand	97	
1		8.1
4		32.3
Steel	480	
¼		10.0
Wood	24–28	
1		2 .4

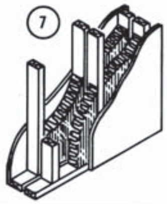
4.3.2 High-Loss Frame Walls

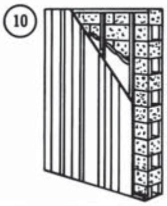
The literature describing high TL walls is extensive. Presented here is a dependable, highly simplified overview of the data with an emphasis on practical solutions for sound room walls. Jones's summary shown in Table 4-7 describes eight frame constructions including the STC performance of each.[7] In each of these constructions Gypsum wallboard is used because it provides an inexpensive and convenient way to get necessary wall mass and as fire retardant properties. Two lightweight concrete block walls, systems 9 and 10, fall in the

Table 4-7. Sound Transmission Class of Some Common Building Partitions*


③


⑤


⑦


⑩

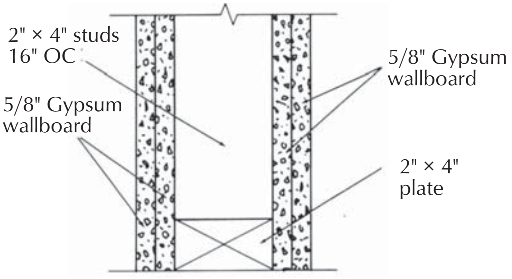
Wall System	Test Sponsor[1]	Laboratory Test Reference	Surface Weight, lb/ft²	STC Rating: No Cavity Absorption	STC Rating: Cavity Absorption
1. Single-row 2 x 4 wood stud (16-inch on center), single-layer $\frac{5}{8}$-in gypsum board panels each side, *direct attached*	OCF	OCF 424 & OCF 423	—	34	36 ($3\frac{1}{2}$-inch glass fiber)
	NGC	NGC 2403 & NGC 2166	6.1	35	38 ($3\frac{1}{2}$-inch glass fiber)
2. *Same as 1*, except double-layer $\frac{1}{2}$-inch gypsum board each side	OCF	OCF W-23-69 & OCF W-25-69	—	39	45 ($3\frac{1}{2}$-inch glass fiber)
3. Single-row 24-gage $3\frac{5}{8}$-inch steel stud (24-inch on center) single $\frac{5}{8}$-in gypsum board panels each side, *direct attached*	NGC	NGC 2385 & NGC 2386	5.2	42	47 ($2\frac{1}{2}$-inch glass fiber)
4. *Same as 3*, except double-layer $\frac{1}{2}$-inch gypsum board each side	NGC	NGC 2282 & NGC 2288	8.9	48	53 (3-inch glass fiber)
5. Single-row 2 x 4 wood stud, single-layer $\frac{5}{8}$-inch gypsum board panels, *direct attached* one side, attached to metal *resilient channels* other side	FPL	TL-73-72	6.4	—	47 ($2\frac{1}{2}$-inch glass fiber)
	OCF	OCF 431 & OCF 427	—	40	46 ($3\frac{1}{2}$-inch glass fiber)
	GA	TL 77-138	—	—	50 ($3\frac{1}{2}$-inch glass fiber)
6. *Same as 5*, except double-layer $\frac{5}{8}$-inch gypsum board each side	USG	TL 67-212 & TL 67-239	10.6	49	59 ($3\frac{1}{2}$-inch mineral fiber)
	NGC	NGC 2368 & NGC 2365	11.3	50	54 ($3\frac{1}{2}$-inch glass fiber)
7. Double-row 2 x 4 wood stud, 1-inch plate separation, single layer $\frac{5}{8}$-inch gypsum board each side	FPL	TL 75-83	7.6	—	57 (double $3\frac{1}{2}$-inch glass fiber)
	OCF	OCF W-43-69 & OCF 448	—	45	56 (3-inch glass fiber)
8. *Same as 7*, except double-layer gypsum board each side	FPL	TL 75-82	12.2	—	63 (double $3\frac{1}{2}$-inch glass fiber)
	OLF	OCF W-42-69 & OCF W-40-69	—	58	62 ($1\frac{1}{2}$-inch glass fiber)
9. 8-inch lightweight hollow concrete block both sides sealed with latex paint	ABPA	TL 70-16	34	46	—
10. *Same as 9*, with addition of furred out wall: $1\frac{5}{8}$-inch 24-gage metal studs, runners placed $\frac{1}{4}$-inch from concrete wall, covered with $\frac{1}{4}$-inch prefinished hardboard facing	ABPA	TL 70-14	36	—	57 ($1\frac{1}{2}$-inch mineral fiber)

[1] OCF—Owens-Corning Fiberglas Corporation
NGC—Gold Bond Building Products Division
FPL—USDA Forest Products Laboratory
GA—Gypsum Association
USG—United States Gypsum
ABPA—American Board Products Association

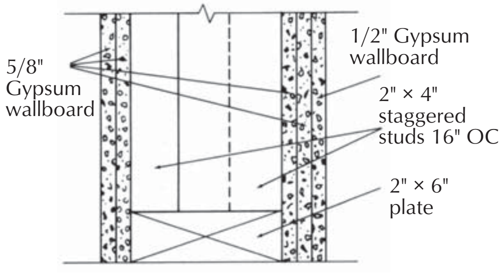
*Source: Reference 9

general STC range of the gypsum wallboard walls 1 to 8, inclusive.

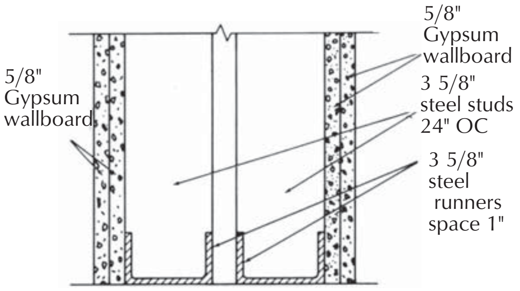
The three papers by Green and Sherry report measurements on many wall configurations utilizing gypsum wallboard.[11] Fig. 4-14 describes three of them yielding STCs from 56 to 62.



A. STC–56



B. STC–58



C. STC–62

Figure 4-14. Three arrangements of gypsum wallboard two-leaf partitions having progressively higher STC ratings. (After Green and Sherry, Reference 11.)

An expression of the empirical mass law stated as an STC rating rather than transmission loss[10] is shown in Fig. 4-15. This makes it easy to evaluate the partitions of Table 4-7 and Fig. 4-14 with respect to partition surface weight. The numbered STC shaded ranges of Fig. 4-15 correspond to the same numbered partitions of Table 4-7, and the A, B, and C points refer to the A, B, and C constructions of Fig. 4-14. From Fig. 4-15 it can be seen that the performance of wall types 1 and 9 can be predicted from the mass law. The other wall types perform better than what the mass law curve predicts. This better performance results primarily from decoupling one leaf of a structure from the other.

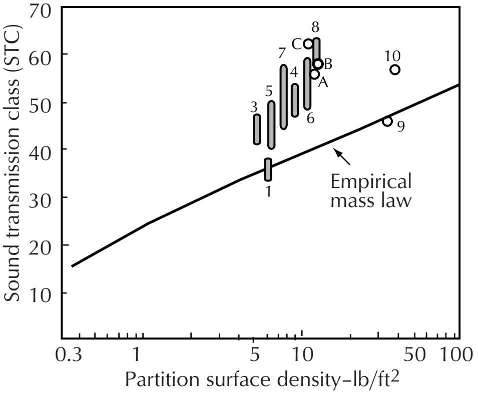


Figure 4-15. A variation of the empirical mass law expressed in terms of sound transmission class rather than TL. The numbers refer to the partitions of Table 4-7; the letters refer to Fig. 4-14.

In recent years there have been new developments in wallboard. QuietRock™*1 is an internally damped wallboard. Although it is considerably more expensive than standard gypsum board, it far outperforms conventional drywall, and for a given STC because less material is needed when using QuietRock, the cost can be offset.

Following are ten points to remember concerning frame walls for highest STC ratings:

1. It is theoretically desirable to avoid having the coincidence dip associated with one leaf of a wall at the same frequency as that of the other leaf. Making the two leaves different with coincidence dips appearing at different frequencies should render their combined effect more favorable. However, Green and Sherry found this effect negligible when partitions having equivalent surface weights were compared.[11]
2. The two leaves of a wall can be made different by utilizing gypsum board of different thickness, mounting a soft fiber (sound-deadening) board under one gypsum board face and/or mounting gypsum board on resilient channels on one side.
3. Resilient channels are more effective on wood studs than on steel studs.

1*QuietRock copyright Quiet Solution, 125 Elko Dr., Sunnyvale, CA, 94089.

4. Steel stud partitions usually have an STC from two to ten points higher than the equivalent wood stud partition. The flange of the common C-shaped steel stud is relatively flexible and transmits somewhat less sound energy from face to face.
5. If multiple layers of gypsum board are used, mounting the second layer with adhesive rather than screws can affect an STC increase by as much as six points. This is especially helpful with higher density walls.
6. A fiberglass cavity filler (such as R-7) may increase STC by five to eight points. It is more effective in multilayer partitions if the second layer is attached with adhesive.
7. A slight increase in STC results from increasing stud spacing from 16 inches to 24 inches on center.
8. Increasing stud size from 2 inches to 3 inches does not significantly increase either transmission loss or STC in steel-stud partitions with filler in the cavity.
9. Additional layers of gypsum wallboard increase STC and TL, but the greatest improvement is with lighter walls. Adding layers increases stiffness, which tends to shift the coincidence dip to a lower frequency.
10. Attaching the first wallboard layer to studs with adhesive actually reduces STC.

4.3.3 Concrete Block Walls

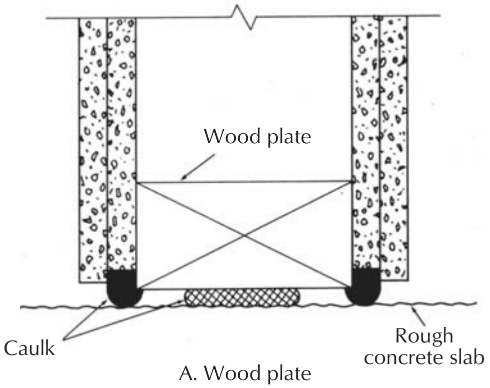
Concrete block walls behave much like solid walls of the same surface weight and bending stiffness. In Table 4-7, wall system 9 is a lightweight, hollow, concrete block wall with both sides sealed with latex paint. In Fig. 4-15 we see that the performance of this specific wall falls close to pure mass operation. The STC-46 is matched or exceeded by many frame walls listed before it in Table 4-7. Wall system 10 in Table 4-7 is the same as 9, except wall system 10 has a new leaf, is furred out, and has mineral fiber added to one side in the cavity. These additions increase the STC from 46 to 57. It should be noted that there are less expensive frame structures that perform just as well. The performance of concrete walls can be improved by increasing the thickness of the wall, by plastering one or both faces, or by filling the voids with sand or well-rodded concrete, all of which increase wall mass. The STC performance of such walls can be estimated from Fig. 4-15 when the pounds-per-square-foot surface density is calculated. To further improve the performance one must add a furred-out facing (such as 10) or adding a second block wall with an air space.

4.3.4 Concrete Walls

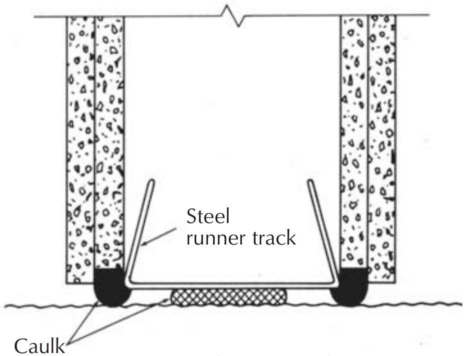
The empirical mass law line in Fig. 4-15 goes to 100 lb/ft^2 (488 kg/m^2), just far enough to describe an 8 inch concrete wall of 150 lb/ft^3 density (surface density 100 lb/ft^2 or 732 kg/m^2). This wall gives a rating close to STC-54. By extending the line we would find that a 12 inch wall would give STC-57, and a concrete wall 24 inches thick, about STC-61. The conclusion is inescapable. This brute-force approach to sound TL is not the cheapest solution. High TL concrete walls can be improved by introducing air space—e.g., two 8 inch walls spaced a foot or so apart. Such a wall requires specialized engineering talent to study damping of the individual leaves of the double wall, the coupling of the two leaves by the air cavity, the critical frequencies involved, the resonances of the air cavity, and so on.

4.3.5 Wall Caulking

There is continual movement of all building components due to wind, temperature expansion and contraction, hygroscopic changes, and deflections due to creep and loading. These movements can open up tiny cracks that are anything but tiny in their ability to negate the effects of a high-loss partition. An acoustical sealant is required to caulk all joints of a partition if the highest TL is to be attained. This type of sealant is a specialty product with nonstaining, nonhardening properties that provides a good seal for many years. Fig. 4-16 calls attention to the importance of bedding steel runners and wood plates in caulking to defeat the irregularities always present on concrete surfaces. A bead of sealant should also be run under the inner layer of gypsum board. The need for such sealing is as important at the juncture of wall-to-wall and wall-to-ceiling as it is at the floor line. The idea is to seal the room hermetically. Fig. 4-17 is a nomograph that illustrates what happens if there is leakage in a partition. The X axis represents a partition that is not compromised by any leaks. The family of curves are gaps or holes expressed as percentages of the whole surface area of the partition. This nomograph shows that a partition rated at a TL of 45 with no penetrations would perform as a TL-30 wall if only 0.1% of the wall were open. Consider what this means in real terms. A partition has a surface area of 10 m^2, 0.1% of 10 m^2 amounts to an opening with an area of a square centimeter (cm^2). This could be a gap in the wall/ floor junction where the caulking was omitted, or it could be the area left open by the installation of an electrical box in a partition. This small gap will reduce the performance of the wall by a significant amount. All of the engineering and calculations that have been discussed so far can be rendered meaningless if sufficient care is not taken to seal *all* holes in a partition.



A. Wood plate



B. Steel runner track

Figure 4-16. Caulking methods used for partitions.

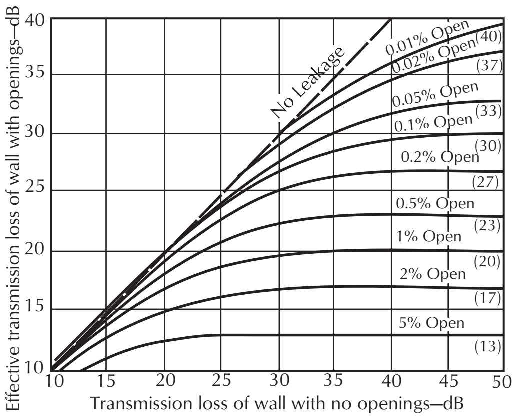
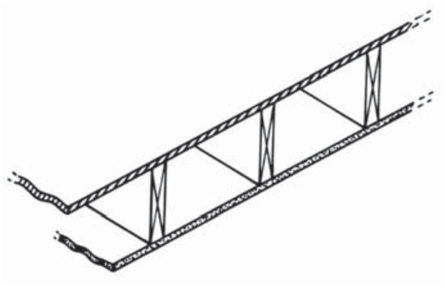
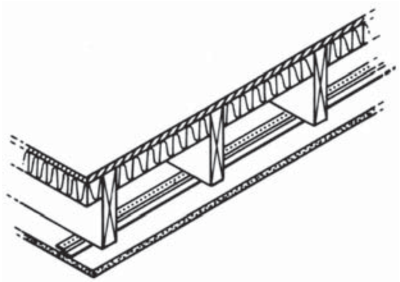


Figure 4-17. Effect of gaps on transmission loss. Courtesy of Russ Berger, Russ Berger Designs.



A. Commonly found construction that passes impact noise from the floor above to the one below with little loss.



B. Similar system having greatly increased transmission loss resulting from the suspension of the ceiling on resilient channels and an absorbent introduced into the air space.

Figure 4-18. Floor and ceiling systems.

4.3.6 Floor and Ceiling Construction

Building high TL walls around a sound room is futile unless similar attention is given to both the floor/ceiling system above the room and to the floor of the sound room itself. Heel and other impact noise on the floor above the room is readily transmitted through the ceiling structure and radiated into the sound room unless precautions are taken. The floor and ceiling structure of Fig. 4-18A is the type common in most existing frame buildings. Impact noise produced on the floor above is transmitted through the joists to the ceiling diaphragm below and radiated with little loss into the room below. Carpet on the floor above softens heel taps, but is low mass, and therefore has little effect on transmission of structure-borne sounds. Some decoupling of the floor membrane from the ceiling membrane is introduced in Fig. 4-18B in the form of resilient mounting of the ceiling gypsum board. Placing absorbent material in the cavity is also of modest benefit. In Table 4-8 four floor and ceiling structures are described along with STC ratings for each, as determined from field TL measurements.

Table 4-8. Floor and Ceiling Systems

Ceiling Treatment	Treatment of Floor Above	Sound Transmission Class*
½ inch gypsum wallboard nailed to joists	1½ inch lightweight concrete on ⅝ inch plywood, 2 × 12 joists 16 inches on centers	STC-48
3 inch mineral wool. Resilient channels 2 ft-0 inch on center ½ inch sound deadening board, ⅝ inch gypsum board	1⅛ inch plywood on 2 × 10 joists 16 inches oc	STC-46
3 inch mineral wool. Resilient Channels 2 ft-0 inch oc ⅝ inch gypsum board	1½ inch lightweight concrete on ½ inch sound deadening board on ⅝ inch plywood, 2 × 10 joists, 16 inches oc	STC-57
2 inch mineral wool. ½ inch sound deadening board. Resilient channels 2 ft-0 in oc ⅝ in gypsum board	1½ inch lightweight concrete on ⅝ inch plywood, 2 × 10 joists, 16 inches oc	STC-57

*These are FSTC ratings.[9]

Another means of decoupling the floor above from the sound room ceiling involves suspending the entire ceiling by a resilient suspension, such as in Fig. 4-19. Mason Industries, Inc. reports one test that demonstrates the efficacy of this approach.[13] They started with a 3 inch concrete floor that alone gave STC-41. With a 12 inch air gap, a ⅝ inch gypsum board ceiling was supported on W30N spring and neoprene hangers, resulting in STC-50. By adding a second layer of ¾ inch gypsum board and a sound-absorbent material in the air space, an estimated STC-55 was realized. The W30N hanger uses both a spring and neoprene. This combination is effective over a wide frequency range. The spring is effective at low frequencies and the neoprene at higher frequencies.

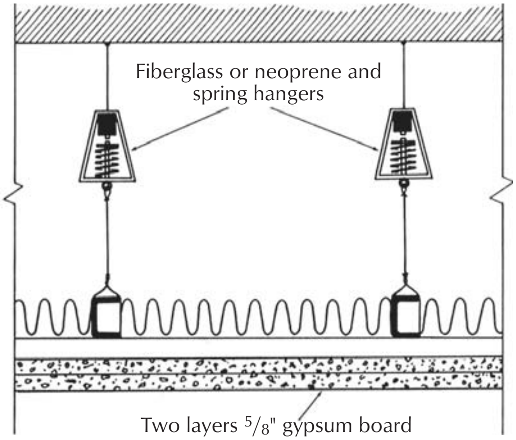

Fiberglass or neoprene and spring hangers
Two layers 5/8" gypsum board

Figure 4-19. A method of suspending a ceiling that gives a great improvement in STC rating of the floor and ceiling combination.

4.3.7 *Floor Construction*

Many variables must be considered when designing isolated floors. These variables include cost, load limits of the existing structure, the desired isolation, and the spectrum of the noise. Every successful system uses a combination of mass and resilient support designed to work above the resonance point of the system, and thus achieve isolation. There are three general approaches to floating or isolated floors; the continuous underlayment, the resilient mount, and the raised slab, Fig. 4-20.

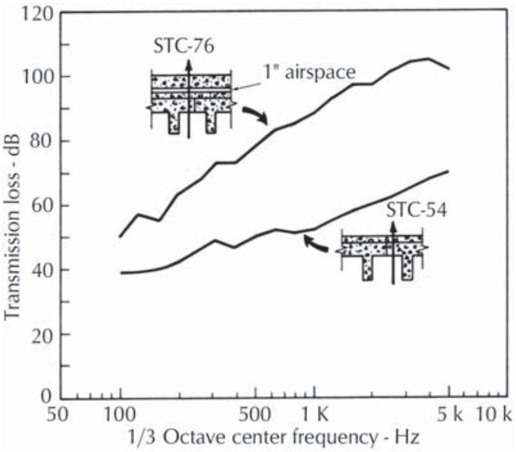
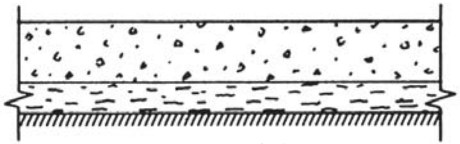
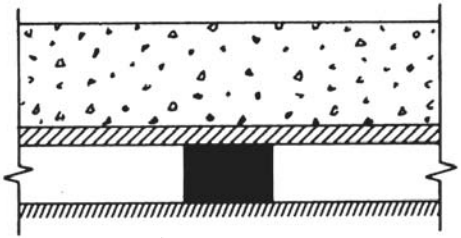


Figure 4-20. The dramatic improvement of sound transmission class (STC) from 54 to 76 by adding a 4 inch floating floor with a 1 inch air gap between it and the structural floor. (Riverbank TL-71-247 test reported by Mason Industries, Inc., in Reference 13.)

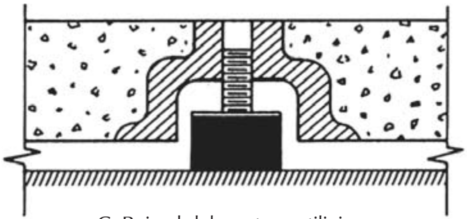
Floating Floors. Once again, simply increasing mass is often the least productive way to make significant gains in STC. For example, a 6 inch solid concrete floor has an STC of 54, and doubling the thickness to 12 inches raises it only to STC-59. There are many recording studios and other sound-sensitive rooms that require floors greater than STC-54. The answer is in dividing available mass and placing an air space between. The results of an actual test, sponsored by Mason Industries, Inc., are given in Fig. 4-21.[12] The TL of basic T sections (4 inch floor thickness) with 2 inches of poured concrete gives a total thickness of 6 inches and the STC-54 mentioned previously. Adding a 4 inch concrete floor on top of the same structural floor with 1 inch of air gap gives a healthy STC-76, which should be adequate for all but the most critical applications. A 4 inch slab added to the 6 inch floor without an air space gives only STC-57. A 19 dB improvement can be attributed directly to the air space.



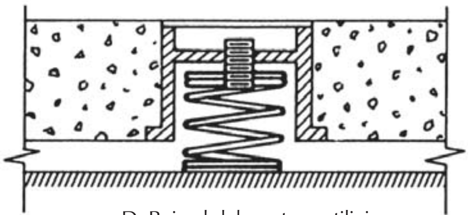
A. Continuous underlayment



B. Isolation mount system



C. Raised-slab system utilizing neoprene for resiliency



D. Raised-slab system utilizing springs for resiliency

Figure 4-21. Four methods used in floating floors for increasing transmission loss.

Continuous Underlayment. The continuous underlayment is the simplest and easiest form of floating floor to construct. It is most often used for residential and light commercial applications where surface loads are relatively light. The technique consists of laying down some sort of vibration-absorbing mat andthen constructing a floor on top of the mat, taking care not to penetrate the mat with any fasteners. The perimeter is surrounded with a perimeter isolation product and sealed with a nonhardening acoustical sealant. Maxxon offers a number of products including Acouti-Mat 3, Acousti-MatII-Green, and Enkasonic®. These are all underlayments that form a resilient layer upon which a wood floor can be constructed, Fig. 4-22, or can be part of a poured concrete system.

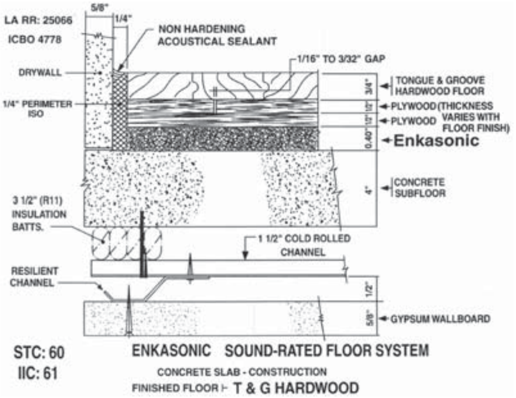

LA RR: 25066
ICBO 4778
5/8"
1/4"
NON HARDENING ACOUSTICAL SEALANT
1/16" TO 3/32" GAP
DRYWALL
1/4" PERIMETER ISO
3/4"
TONGUE & GROOVE HARDWOOD FLOOR
1/2"
PLYWOOD (THICKNESS
1/2"
PLYWOOD VARIES WITH FLOOR FINISH)
0.40"
Enkasonic
4"
CONCRETE SUBFLOOR
3 1/2" (R11) INSULATION BATTS.
1 1/2" COLD ROLLED CHANNEL
RESILIENT CHANNEL
1/2"
5/8"
GYPSUM WALLBOARD
STC: 60
IIC: 61
ENKASONIC SOUND-RATED FLOOR SYSTEM
CONCRETE SLAB - CONSTRUCTION
FINISHED FLOOR ⊢ T & G HARDWOOD

Figure 4-22. Enkasonic floor system.

Isolation Mount Systems. If heavier loads are anticipated and greater isolation is needed, an isolation mount system should be considered. Various manufacturers build systems for isolating either wood floors or concrete slabs. Wood floors can be isolated as shown in Fig. 4-23. This system offered by Kinetics utilizes encapsulated fiberglass pads, imbedded in a roll of low-frequency fiberglass designed to fill the air space.

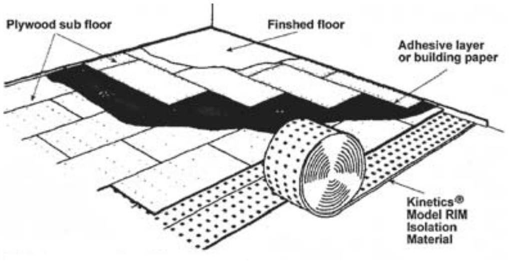

Plywood sub floor
Finshed floor
Adhesive layer or building paper
Kinetics® Model RIM Isolation Material

Figure 4-23. Kinetics Floating Wood Floor. Courtesy Kinetics Corp.

Another approach by Mason Industries is to build a grid supported on neoprene mounts or, if greater isolation is needed, on combination spring and neoprene as shown in Fig. 4-24. A wood floor is then built on the substructure.

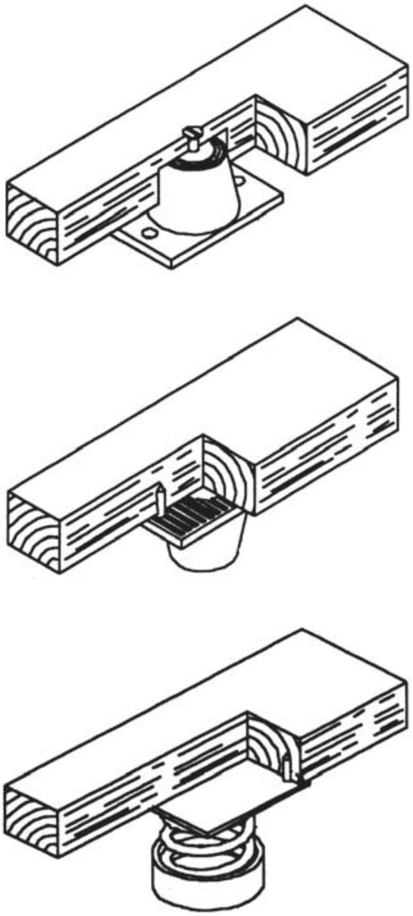


Figure 4-24. Mason Industries Floating Wood Floor systems using either springs or neoprene. Courtesy Mason Industries.

In some situations a floating concrete slab is indicated in Fig. 4-25, concrete slab is supported by the model RIM mat. The roll-out mat is ordered with the pad spacing based on the expected load. When the mat is unrolled (1) the plywood panels are then put in place, (2) the plastic sheet laid over the plywood, and (3) the concrete poured, Fig. 4-25. A perimeter board isolates the floating floor from the walls. The plastic film protects the plywood and helps to avoid *bridges*.

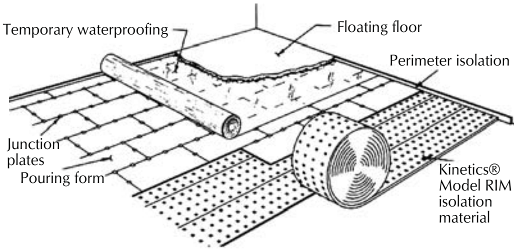

Temporary waterproofing
Floating floor
Perimeter isolation
Junction plates
Pouring form
Kinetics® Model RIM isolation material

Figure 4-25. The roll-out mat system of constructing floating floors. Courtesy Kinetics Inc.

Raised-Slab or Jack-Up System. This system is for heavy duty applications where high STC ratings are needed. In Fig. 4-26 the individual isolators are housed in metal canisters, Fig. 4-27, that are placed typically on 36 inch to 48 inch centers each dimension. The metal canisters are arranged to tie into the steel reinforcing grid and are cast directly in the concrete slab. After sufficient curing time (about 28 days), it is lifted by judicious turning of all the screws one-quarter or if

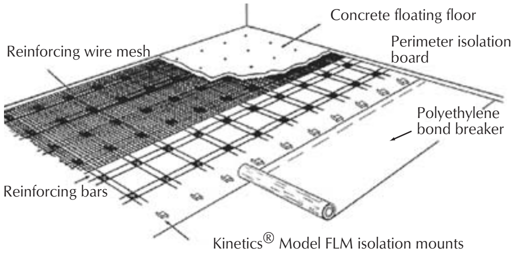

Concrete floating floor
Reinforcing wire mesh
Perimeter isolation board
Polyethylene bond breaker
Reinforcing bars
Kinetics® Model FLM isolation mounts

Figure 4-26. Kinetics FLM Jack Up Concrete Floor system. Courtesy Kinetics Corp.

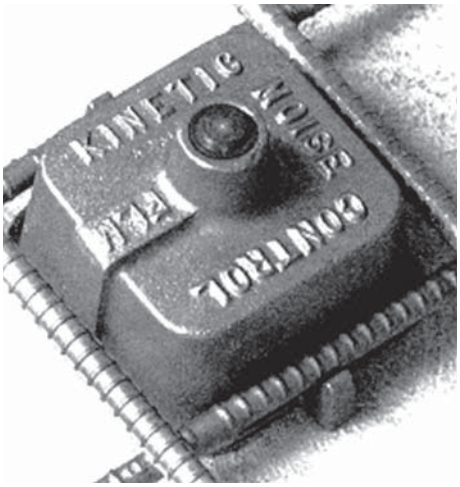

KINETIC
NOISE
CONTROL

Figure 4-27. Kinetics FLM isolation mount. Courtesy Kinetics Corp.

one-half turn at a time. This is continued until an air space of at least 1 inch is achieved. Fig. 4-28 shows an alternative raised slab system utilizing springs instead of neoprene or fiberglass mounts. After the slab is raised to the desired height, the screw holes are filled with grout and smoothed. Fig. 4-29 further describes the elements of the raised-slab system. Turning the screws in the load-bearing isolation mounts raises the cured slab, producing an air space of the required height. This system requires heavier reinforcement rods in the concrete than the system of Fig. 4-25.

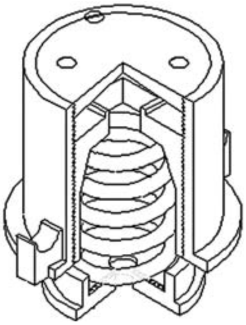


Figure 4-28. Mason Industries FS spring jack up floor system. Courtesy Mason Industries.

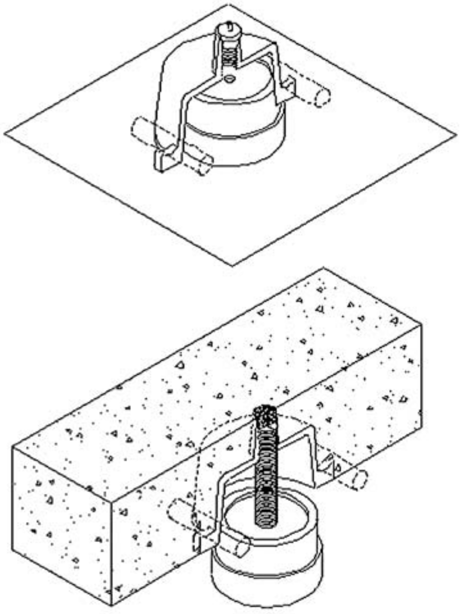


Figure 4-29. Details of a jack mount. Courtesy Mason Industries.

4.3.8 Summary of Floating Floor Systems

Loading must be calculated for each type of floating floor systems discussed. If the resilient system is too stiff, vibration will travel through the isolator rendering it ineffective. Likewise if the springs are too soft, they will collapse under the weight of the structure and also be ineffective.

Each floating floor system has its advocates. No one type of floor will suit all situations. The designer is urged to consider all the variables before making a decision. For example, there are pros and cons concerning use of neoprene versus the compressed, bonded, and encased units of glass fiber. Most of the arguments have to do with deterioration of isolating ability with age and freedom from oxidation, moisture penetration, and so on.

Fig. 4-30 combines several features that have been discussed in a "room within a room." The walls are supported on the floating floor and stabilized with sway braces properly isolated. The ceiling is supported from the structure with isolation hangers. This type of hanger incorporates both a spring, which is particularly good for isolation from low-frequency vibration, and a Neoprene or a fiberglass element in series, which provides good isolation from higher-frequency components. An important factor is the application of a non-hardening type of acoustical sealant at the points marked "S." An even better approach would be to support the ceiling from the walls by using joists or trusses spanning the room. Such a room should provide adequate protection from structure-borne vibrations originating within the building as well as from those vibrations transmitted through the ground to the building from nearby truck, surface railroad, or subway sources.

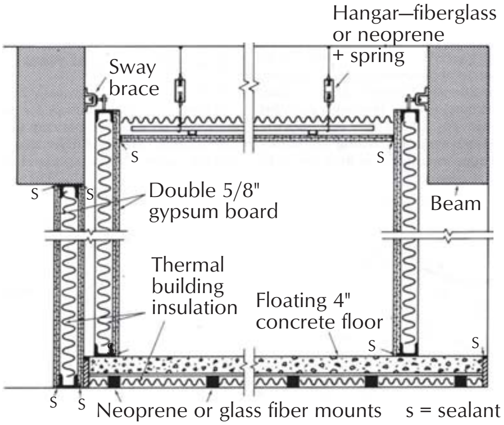

Hangar—fiberglass or neoprene + spring
Sway brace
S
S
S
S
Double 5/8" gypsum board
Beam
Thermal building insulation
Floating 4" concrete floor
S
S
S
S
Neoprene or glass fiber mounts
s = sealant

Figure 4-30. A "room-within-a-room" exemplifying the principles discussed in the text.

The design of rooms to achieve maximum isolation from airborne and structure-borne sounds is a highly specialized undertaking, ordinarily entrusted to consultants expert in that branch of acoustics. However, a sound engineer, charged with the responsibility of working with a consultant or doing the design personally, is advised to become familiar with the sometimes conflicting claims of suppliers and the literature on the subject.

4.3.9 Acoustical Doors

Every part of an acoustical door is critical to its performance. Special metal acoustical doors are available with special cores, heavy hinges, including special sealing and latching hardware. Their acoustical performance is excellent and their higher cost must be evaluated against high labor costs in constructing an alternative. There are two design elements required in considering what kind of door to utilize. There is the transmission loss of the door itself and there is the sealing system. The sealing system is the more critical of the two. Whatever system is used, it must hold up over time and withstand the wear and tear of use. Doors and their seals are difficult to build and are often the weak point of a sound room. There is good reason to design sound room access and egress in such a way that excessively high performance is not required of a single door. Use of a sound lock corridor principle places two widely spaced doors in series, relieving the acoustical requirements of each, Fig. 4-31.

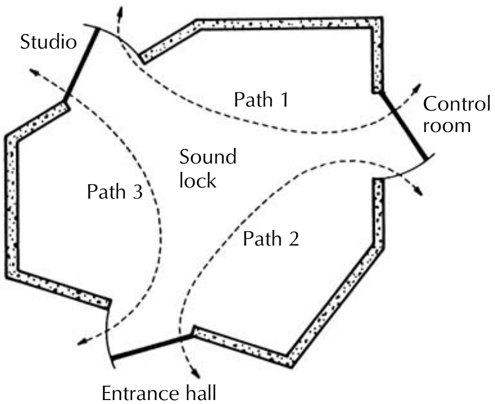


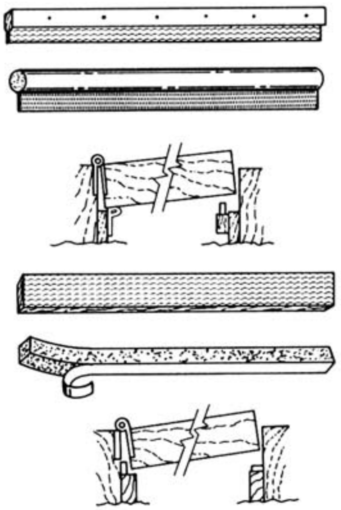
Figure 4-31. Sound lock corridor.

Homemade Acoustical Doors. An inexpensive door, satisfactory for less demanding applications, can be built from void-free plywood or high density particle board. It is also possible to start with a core material of particle board and laminate it with gypsum board if sufficient care is taken to protect the fragile edges of the gypsum board. Doors for acoustical isolation must have a solid and void-free core and be as massive as practical. Most residential grade doors are hollow and approach acoustical transparency. Some commercially available solid core doors are made of laminated wood; others, of particle board with composition board facing. The latter has the greater surface density. The 5.2 lb/ft^2 of the particle-board type gives an STC value of about 35. An STC-35 does not do justice to, say, STC-55 walls. Nevertheless, for doors separated as they are in the case of a sound lock, the TL of one door comes close to adding arithmetically to the loss of the other door. Two doors, well separated, approach doubling the effect of one.

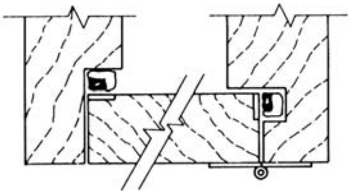
All this implies a perfect seal around the periphery of the door attained only by nailing the door shut and applying a generous bead of acoustical sealant on the crack. A practical operative door must utilize some form of weatherstripping or other means for its seal. Fig. 4-32 illustrates different approaches to sealing a door.[13] Many of these, especially the wiping type, require constant maintenance and frequent replacement. One of the more satisfactory types is the magnetic seal, similar to those on most household refrigerator doors. Zero International manufactures a system of door seals specifically designed for acoustical applications, Fig. 4-33. This type of commercially available acoustical door seal is a good way to get results from a homemade door that approaches the performance of a proprietary door at a fraction of the cost.

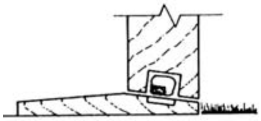
Proprietary Acoustical Doors. By far the more satisfactory doors for acoustical isolation in sound rooms are those manufactured especially for the purpose. Such doors offer measured and guaranteed performance over the life of the door with only occasional adjustment of seals. This is in stark contrast to the need for constant seal maintenance in the homemade door shown in Fig. 4-32. Each manufacturer has its own strengths. Some doors like the Overly and the IAC use cam lift hinges, which actually lift the door as it opens.

Manufacturers of building elements that need to be rated for sound transmission use ASTM standards in measuring their products. ASTM e-90 is the appropriate standard for sound transmission measurements. Copies of the standards are available at www.ASTM.com. Most manufacturers build a range of doors to suit specific needs. IAC builds doors ranging from an STC-43 to an impressive STC-64, Fig. 4-34.

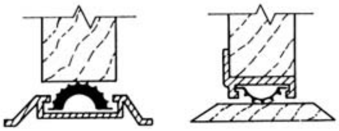


A. Foam or felt strips.

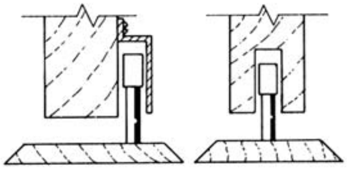




B. Magnetic.



C. Wiping rubber seal.



D. Drop bar.

Figure 4-32. Numerous types of weather stripping can be used for sealing doors to audio rooms. Courtesy Tab Books, Inc.

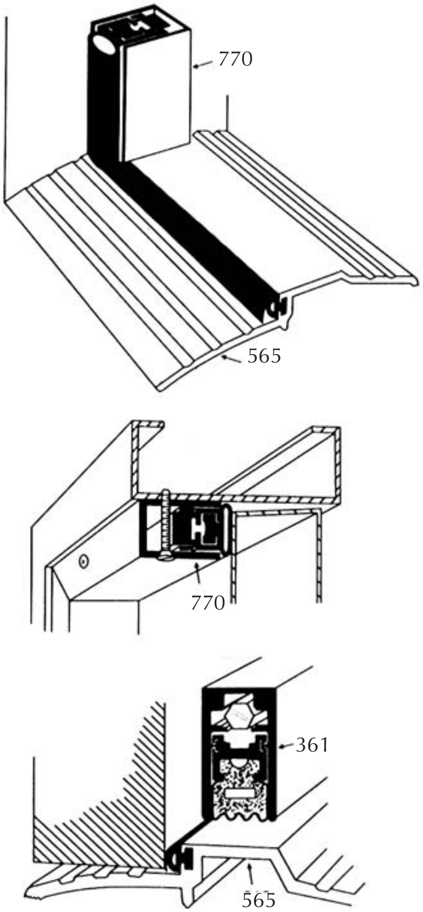


Figure 4-33. Sealing systems from Zero Mfg.

4.3.10 Windows

Occasionally sound rooms require windows. The observation window between control room and studio is an example. It can very easily have a weakening effect on the overall TL of the partition between the two rooms. (See Section 3.3.11.) A wall with a rating of STC-60 alone might very well be reduced to STC-50 with even one of the more carefully designed and built windows installed. Just how much the window degrades the overall TL depends on the original loss of the partition, the TL of the window alone, the relative areas of the two, and of course, the care with which the window is installed. To understand the factors going into the design of an effective observation window, a good place to start is to study the effectiveness of glass as a barrier.[15]

Transmission Loss of Single Glass Plates. The measured transmission loss of ¼ inin, ½ inch, and ¾ inch single-glass plates (or float) 52 inch × 76 inch is shown in Fig. 4-35. As expected, the thicker the glass plate, the higher the general TL except for a coincidence dip in

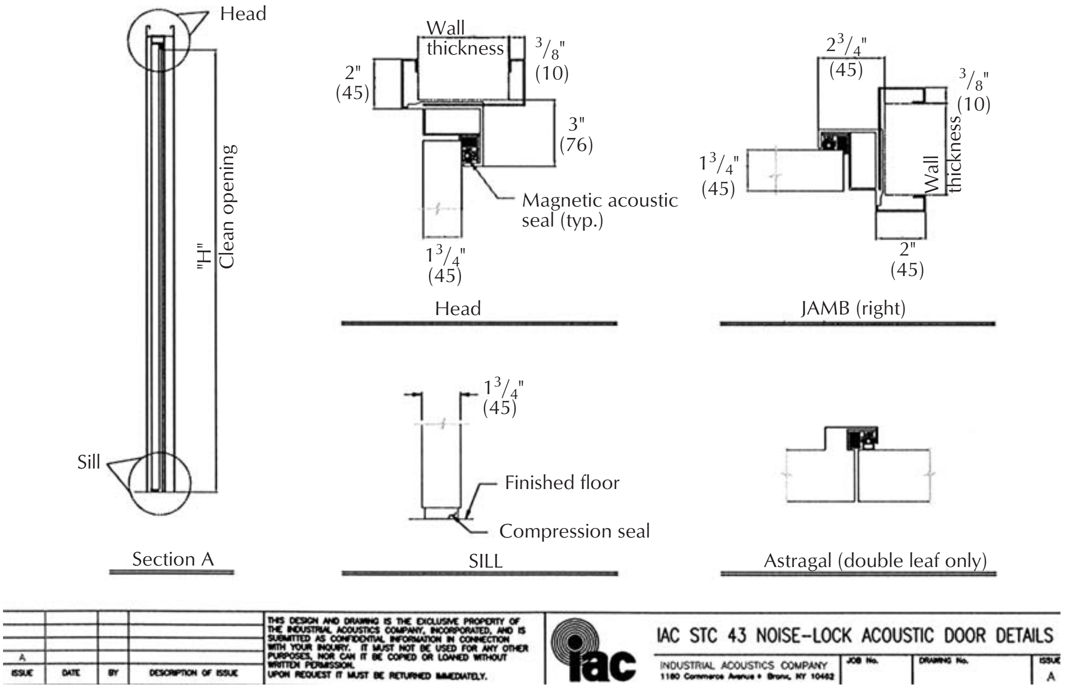

Head
"H"
Clean opening
Sill
Section A
Wall
thickness
2"
(45)
3/8"
(10)
3"
(76)
Magnetic acoustic
seal (typ.)
1 3/4"
(45)
Head
2 3/4"
(45)
3/8"
(10)
1 3/4"
(45)
Wall
thickness
2"
(45)
JAMB (right)
1 3/4"
(45)
Finished floor
Compression seal
SILL
Astragal (double leaf only)
A
ISSUE
DATE
BY
DESCRIPTION OF ISSUE
THIS DESIGN AND DRAWING IS THE EXCLUSIVE PROPERTY OF THE INDUSTRIAL ACOUSTICS COMPANY, INCORPORATED, AND IS SUBMITTED AS CONFIDENTIAL INFORMATION IN CONNECTION WITH YOUR INQUIRY. IT MUST NOT BE USED FOR ANY OTHER PURPOSES, NOR CAN IT BE COPIED OR LOANED WITHOUT WRITTEN PERMISSION. UPON REQUEST IT MUST BE RETURNED IMMEDIATELY.
iac
IAC STC 43 NOISE-LOCK ACOUSTIC DOOR DETAILS
INDUSTRIAL ACOUSTICS COMPANY
1160 Commerce Avenue • Bronx, NY 10462
JOB No.
DRAWING No.
ISSUE
A

Figure 4-34. IAC STC 43 Door. Courtesy IAC.

each graph. Although the heavy ¾ inch plate attains a TL of 40 dB or more above 2 kHz. It is inappropriate for use in an STC-50 or STC-55 wall. Considering this general lack of sufficient TL and the complication of the coincidence dip, the single-glass approach is insufficient for most observation window needs. Laminated glass is more of a *limp mass* than glass plate of the same thickness and, hence, has certain acoustical advantages in observation windows. The characteristics of ¼ inch, ½ inch, and ¾ inch laminated single-glass plates are shown in Fig. 4-36.

Transmission Loss of Spaced Glass Plates. Fig. 4-37 shows the effect of three different spacings. In all cases the same ½ inch and ¼ inch glass plates are used, but the air space is varied from 2 inches to 6 inches. The effect of spacing the glass plates is greatest below 1500 Hz. There is practically no increase in transmission loss by spacing the two glass plates above 1500 Hz. In general, the 2 inch increase from 2 inches to 4 inches is less effective than the same 2 inch spacing increase from 4 inches to 6 inches. Many observation windows in recording studios utilize spacings of 12 inches or more to maximize the spacing effect.

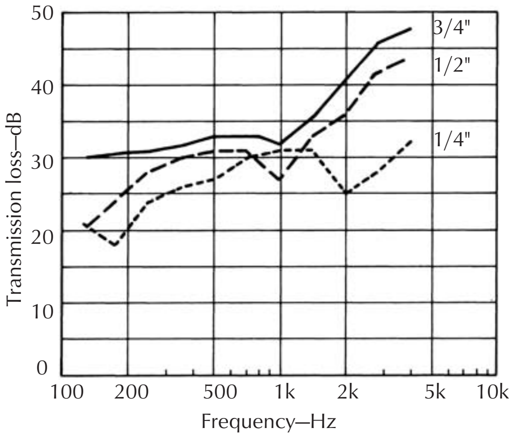


Figure 4-35. Sound TL characteristics of single glass (plate or float) panels. Courtesy Libbey-Owens-Ford Co. (After Reference 16.

When two glass plates are separated only a small amount, such as glass widely used for heat insulation, the sound TL is essentially the same as the glass alone from which it is fabricated. There is practically no acoustical advantage using this type of glass in observation windows. This is one of the few cases where thermal insulation does not correspond to acoustic isolation. The single case of using laminated glass for one of

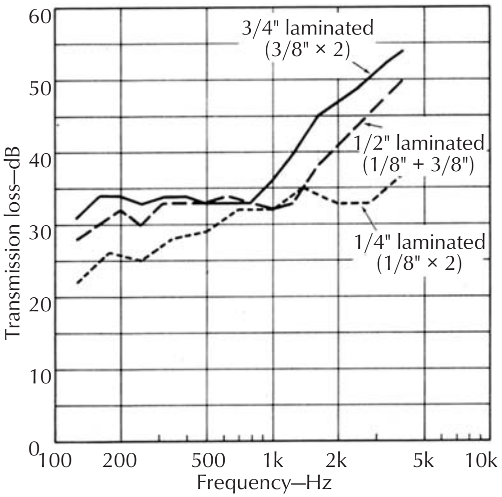


Figure 4-36. Sound TL characteristics of single panels of laminated glass. Courtesy Libbey- Owens-Ford Co. (After Reference16.)

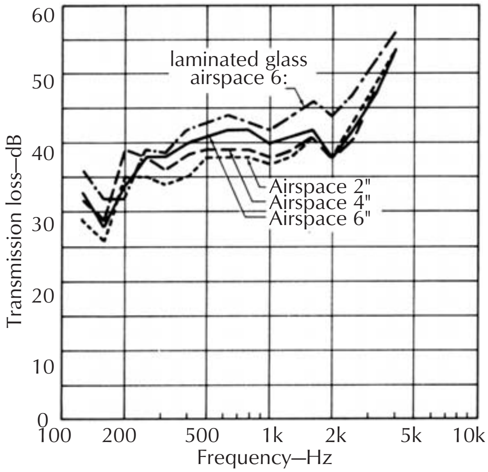


Figure 4-37. Spacing two dissimilar glass plates improves transmission loss. Glass of ½ and ¼ inch thickness used in all cases.

the plates with 6 inch separation is included in Fig. 4-37. The superior performance of laminated glass comes with a higher cost.

Managing Cavity Resonance. The TL measurements in Fig. 4-37 were made with no absorbing material around the periphery of the space between the two glass plates. By lining this periphery with absorbent material, the natural cavity resonance of the space is reduced. An average 5 dB increase in TL can be achieved by installing a minimum of 1 inch absorbent on these *reveals*. The use of 4 inches of absorbing material, covered with, perhaps, perforated metal, further improves low-frequency transmission loss.

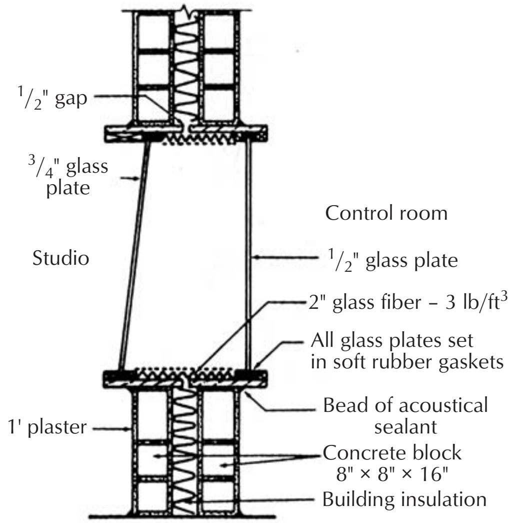
The practice of using glass plates of different thickness is substantiated by shallower coincidence dips in Fig. 4-37 as compared to Fig. 4-35. Resonance associated with the plates or the cavity tend toward the creation of *acoustical holes*, or the reduction of TL at the resonance frequencies. Hence, distributing these resonance frequencies by the staggering of plate thickness and use of laminated glass is important.

Homemade Acoustical Windows. The essential constructional features of two types of observation windows are shown in Fig. 4-38. Fig. 4-38A is typical of the high TL type commensurate with walls designed for high loss. The high TL of the window is achieved by using heavy laminated glass, maximum practical spacing of the glass plates, absorbent reveals between the glass plates, and other important details such as a generous application of acoustical sealant. It is very important to note that the windowsill and other elements of the frame do not bridge the gap between the two walls and thereby compromise the double wall construction. Bridging the double wall construction at the window is a very common error that must be avoided if the STC of the partition is to be maintained.

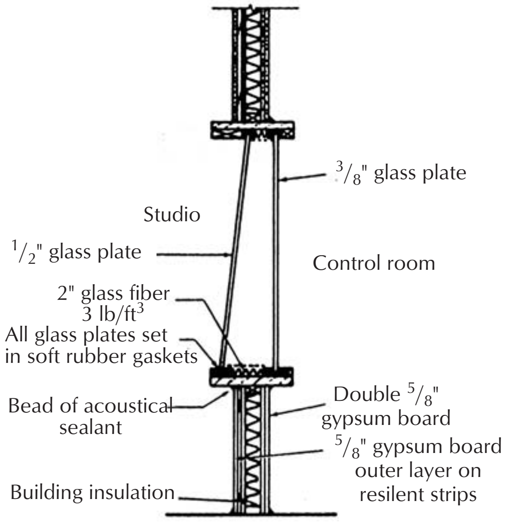
Fig. 4-38B shows a window for a single stud wall, a more modest TL. The same general demands are placed on this window as on the one in Fig. 4-38A, except that scaled down glass thickness and spacing are appropriate.

Inclining one of the plates, as shown in Fig. 4-38, has advantages and disadvantages. Slanting one pane reduces the average spacing, which slightly reduces the TL. However, slanting one window as shown especially in a studio (as distinct from a control room) will have the beneficial effect of preventing a discrete reflection right back at a performer standing in front of the window. The principal benefit of such plate inclination is really the control of *light* reflections that interfere with visual contact between the rooms.

Proprietary Acoustical Windows. Many of the same companies that build proprietary acoustical doors also build acoustical windows. IAC builds a line of windows ranging from STC-35 to STC-58. The STC-53 window from IAC is shown in Fig. 4-39 and Fig. 4-40. It should be noted that the same warning about bridging a double wall construction applies to proprietary windows as well as to home-made ones.


1/2" gap
3/4" glass plate
Control room
Studio
1/2" glass plate
2" glass fiber – 3 lb/ft3
All glass plates set in soft rubber gaskets
Bead of acoustical sealant
1' plaster
Concrete block 8" × 8" × 16"
Building insulation

A. A window suitable for a high TL wall.


3/8" glass plate
Studio
1/2" glass plate
Control room
2" glass fiber 3 lb/ft3
All glass plates set in soft rubber gaskets
Bead of acoustical sealant
Double 5/8" gypsum board
5/8" gypsum board outer layer on resilent strips
Building insulation

B. A window suitable for a more modest frame wall.

Figure 4-38. Construction details for practical observation windows set in a partition between control room and studio.

4.3.11 *Transmission Loss of a Compound Barrier*

We are using the term *compound* to refer to those partitions that are not homogeneous—e.g. those partitions that include areas with differing TL ratings. For example, when an observation window having one TL is set in a wall having another TL, the overall TL is obviously something else, but what is it? It most certainly cannot be obtained by simple manipulation of TLs or STC values. The problem must be referred to as the basics of sound power transmission. Fig. 4-41 illustrates the case of a 4.4 ft × 6.4 ft window set in a 10 ft × 15 ft partition between control room and studio. The way the transmission loss of the window and the wall affect each other is given by the expression:[18,19]

$$TL = -10\log\left(\frac{S_1}{10^{\frac{TL_1}{10}}} + \frac{S_2}{10^{\frac{TL_2}{10}}}\right) \qquad (4\text{-}4)$$

where,

TL is the overall transmission loss,

S_1 is the fractional wall surface,

TL_1 is the wall transmission loss in decibels,

S_2 is the fractional window surface,

TL_2 is the window transmission loss in decibels.

As an example let us say that for a given frequency the wall $TL_1 = 50$ dB and the window $TL_2 = 40$ dB. From Fig. 4-40 we see that $S_1 = 0.812$ and $S_2 = 0.188$. The overall TL is

$$\begin{aligned} TL &= -10\log\left(\frac{S_1}{10^{\frac{TL_1}{10}}} + \frac{S_2}{10^{\frac{TL_2}{10}}}\right) \\ &= 45.7dB \end{aligned}$$

The 40 dB window has reduced the 50 dB wall to a 45.7 dB overall effectiveness as a barrier. This is for a given frequency. Fig. 4-42 solves Eq. 4-4 in a graphical form using the following steps:

1. Figure the ratio of glass area to total wall area, and find the number on the X axis.
2. Subtract window TL from wall TL, and find the intersection of this value with the area ratio on the X axis.
3. From the intersection, find the reduction of the wall TL from the left scale.
4. Subtract this figure from the original wall TL.

Using the graph of Fig. 4-42, find the effect of the window on the compound wall. The ratio of the window area to the wall area is 0.23. Locate 0.23 along the bottom axis. The difference in TL between the two is 10 dB. Find the intersection between the 10 dB line and the ratio of the areas. A reduction of slightly less than 5 dB is read off the left scale. Subtracting 5 dB from the

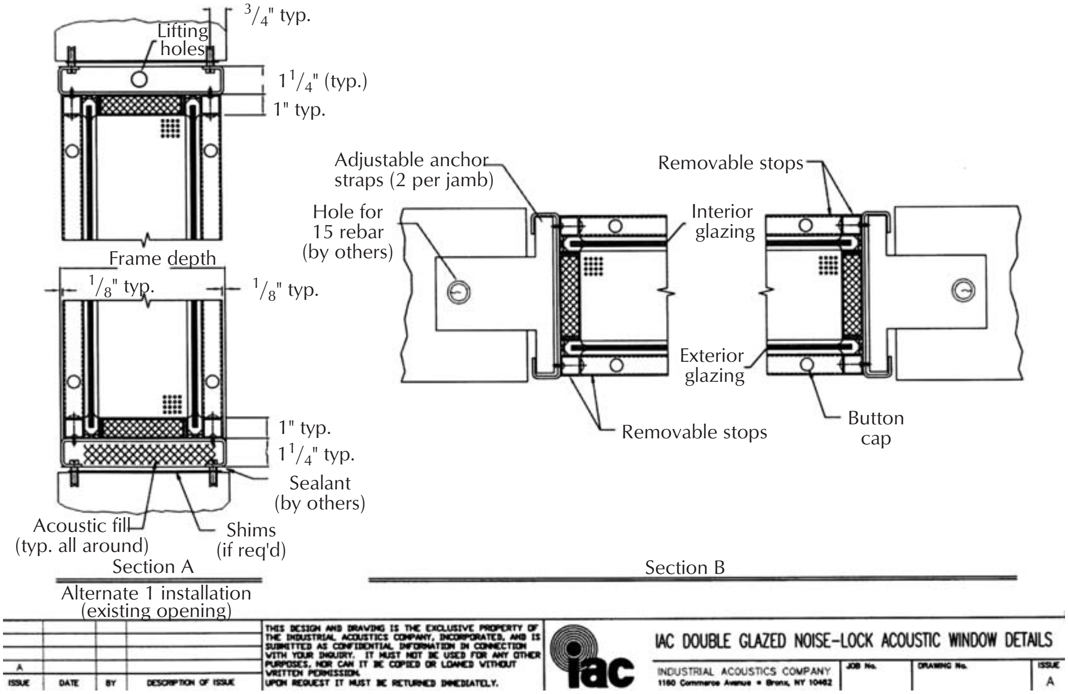

3/4" typ.
Lifting holes
1 1/4" (typ.)
1" typ.
Frame depth
1/8" typ.
1/8" typ.
1" typ.
1 1/4" typ.
Sealant (by others)
Acoustic fill (typ. all around)
Shims (if req'd)
Section A
Alternate 1 installation (existing opening)
Adjustable anchor straps (2 per jamb)
Hole for 15 rebar (by others)
Removable stops
Interior glazing
Exterior glazing
Removable stops
Button cap
Section B
A
ISSUE
DATE
BY
DESCRIPTION OF ISSUE
THIS DESIGN AND DRAWING IS THE EXCLUSIVE PROPERTY OF THE INDUSTRIAL ACOUSTICS COMPANY, INCORPORATED, AND IS SUBMITTED AS CONFIDENTIAL INFORMATION IN CONNECTION WITH YOUR INQUIRY. IT MUST NOT BE USED FOR ANY OTHER PURPOSES, NOR CAN IT BE COPIED OR LOANED WITHOUT WRITTEN PERMISSION. UPON REQUEST IT MUST BE RETURNED IMMEDIATELY.
iac
IAC DOUBLE GLAZED NOISE-LOCK ACOUSTIC WINDOW DETAILS
INDUSTRIAL ACOUSTICS COMPANY
1160 Commerce Avenue • Bronx, NY 10462
JOB No.
DRAWING No.
ISSUE
A

Figure 4-39. Noise-Lock™ window by IAC is rated at STC 53. Courtesy IAC.

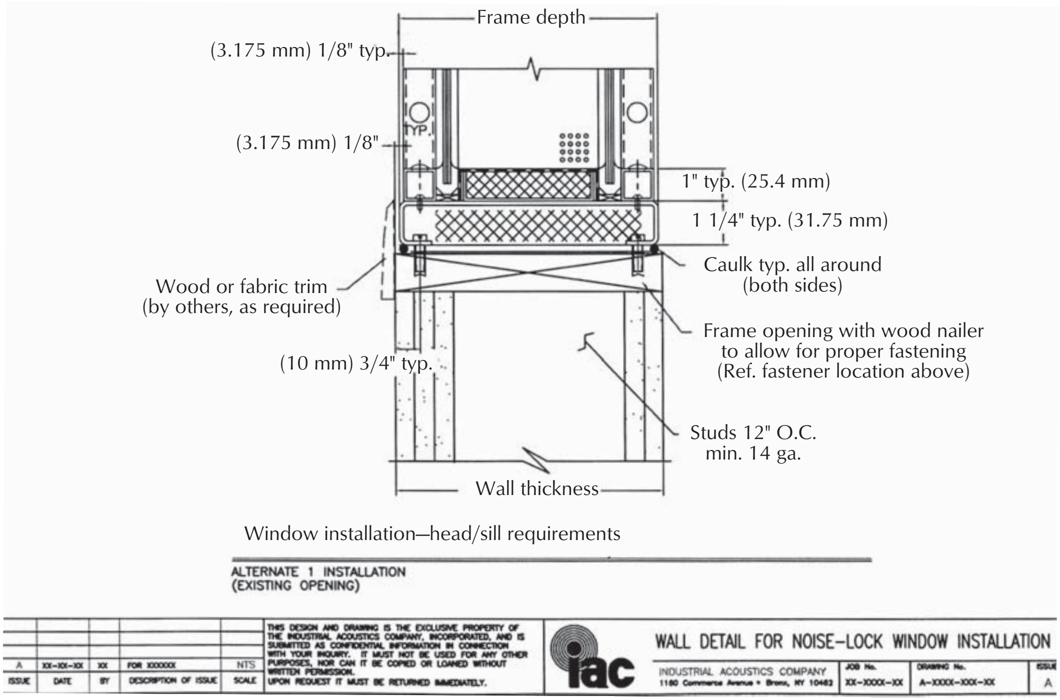

Frame depth
(3.175 mm) 1/8" typ.
(3.175 mm) 1/8"
TYP.
1" typ. (25.4 mm)
1 1/4" typ. (31.75 mm)
Caulk typ. all around (both sides)
Wood or fabric trim (by others, as required)
Frame opening with wood nailer to allow for proper fastening (Ref. fastener location above)
(10 mm) 3/4" typ.
Studs 12" O.C. min. 14 ga.
Wall thickness
Window installation—head/sill requirements
ALTERNATE 1 INSTALLATION (EXISTING OPENING)
A
XX-XX-XX
XX
FOR XXXXXX
NTS
ISSUE
DATE
BY
DESCRIPTION OF ISSUE
SCALE
THIS DESIGN AND DRAWING IS THE EXCLUSIVE PROPERTY OF THE INDUSTRIAL ACOUSTICS COMPANY, INCORPORATED, AND IS SUBMITTED AS CONFIDENTIAL INFORMATION IN CONNECTION WITH YOUR INQUIRY. IT MUST NOT BE USED FOR ANY OTHER PURPOSES, NOR CAN IT BE COPIED OR LOANED WITHOUT WRITTEN PERMISSION. UPON REQUEST IT MUST BE RETURNED IMMEDIATELY.
iac
WALL DETAIL FOR NOISE-LOCK WINDOW INSTALLATION
INDUSTRIAL ACOUSTICS COMPANY
1160 Commerce Avenue • Bronx, NY 10462
JOB No.
XX-XXXX-XX
DRAWING No.
A-XXXX-XXX-XX
ISSUE
A

Figure 4-40. Head and sill requirements for IAC Noise-Lock™ windows. Courtesy IAC.

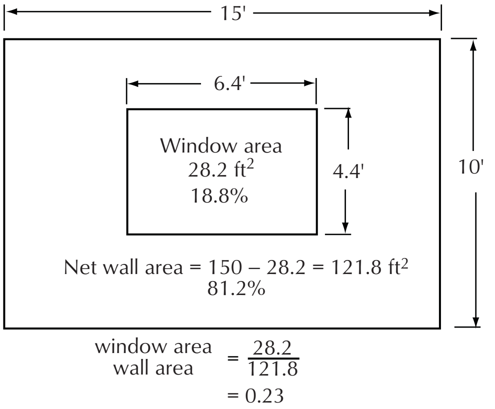

15'
6.4'
Window area
28.2 ft^2
18.8%
4.4'
10'
Net wall area = 150 – 28.2 = 121.8 ft^2
81.2%
$\frac{\text{window area}}{\text{wall area}} = \frac{28.2}{121.8}$
= 0.23

Figure 4-41. Typical observation window in a wall between control room and studio.

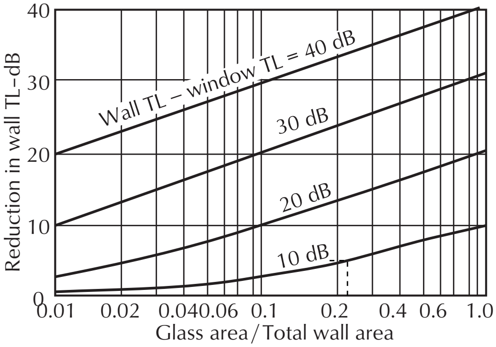


Figure 4-42. Graphical determination of the effect on the overall transmission loss (TL) of a wall by an observation window.

50 dB wall *TL* gives the overall TL with a window of 45 dB. (Calculated from Eq. 4-4 gives 45.7 dB.)

It is usually easier and more economical to get high TL in wall construction than in window construction. The possibility arises of compensating for a deficient window by overdesigning the wall. For example, recognizing that an STC-70 masonry wall is possible, how far will it lift an STC-45 window? Using Eq. 4-4 again, we find the overall STC to be 52.2 dB, an increase of over 7 dB over the STC-45 window. Actually, using Eq. 4-4 with STC values is a gross oversimplification embracing all the inaccuracies of fitting measured TL values with a single-number STC rating. Making the calculations from measured values of TL at each frequency point is much preferred. Of course, all this assumes an airtight seal has been achieved.

Everything that bridges the isolation system is a potential *short circuit* for noise. Such bridges include, HVAC ducts, electrical conduit, sprinkler systems, plumbing, raceways, and the like.

Now we have the formula for empirically looking at the effect of a crack in the wall (Fig. 4-17 was plotted using Eq. 4-4). Let us assume that an observation window and wall combination have a calculated composite TL of 50 dB. The window, installed with less than ideal craftsmanship, developed a ⅛ inch (0.125 in) crack around the window frame as the mortar dried and pulled from the frame. Since this is the window of Fig. 4-41, the length of the crack is 21.6 ft, giving a crack area of 0.225 ft^2. What effect will this crack have on the otherwise 50 dB wall? Substituting into Eq. 4-4, we find the new TL of the wall with the crack to be 28 dB. This is similar to leaving off an entire pane of glass or a layer of gypsum board. If the crack were only 1⁄16 inch wide, the TL of the wall would be reduced from 50 dB to 31.2 dB. A crack only 0.001 inch wide would reduce the TL of 50 dB to 40.3 dB. Let the builder beware!

4.3.12 Isolation Systems Summary

Noise migrates from one area to another in two ways. It travels through the air and it travels through the structure. To reduce or eliminate airborne noise, one must eliminate all air paths between the spaces. To reduce structure-borne noise one must create isolation systems that eliminate mechanical connections between spaces. It is a rather simple matter to make theses statements. Implementing the solutions is obviously much more difficult. The following points should be kept in mind:

- Make seams airtight.
- Analyze all possible flanking paths that noise will take and realize that *all* must be controlled if significant isolation is desired.
- A room built entirely on a floating slab with the ceiling supported entirely by the walls will always be superior to any other method.

4.4 Heating, Ventilating, and Air Conditioning (HVAC) Systems for Low Noise

So far in this chapter we have considered systems that keep unwanted sound out. When we consider HVAC systems we are dealing with systems that (a) breach the acoustical shell designed to keep noise out, (b) introduce considerable noise of their own, and (c) provide a pathway for sound (noise) to easily migrate from one space to another. HVAC systems can sometimes undermine all the efforts of isolation. Often the cheapest solution to providing HVAC to sound sensitive spaces is to use window units that get shut off when quiet is needed! If this solution is not acceptable, and central distributed systems must be used, the designer must understand that success will require significant expense and engineering. The design of HVAC systems is best left to professional mechanical engineers. No better preparation for this responsibility can be obtained than from carefully studying the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE) publications.[16,17,18]

It is important to understand that HVAC systems found in most residences or even in light commercial or office spaces are totally inadequate for use in noise critical spaces. Unlike residential systems that often use high efficiency systems that deliver low volumes of cold air at high velocities, low noise systems require high volume, low velocity delivery. Many commercial systems utilize supply ducts and the return relies on leakage under doors or common ceiling plenums. In order to achieve low noise, both the supply and return must be individually ducted to each room.

4.4.1 Location of HVAC Equipment

From the standpoint of sound room noise, the best location for the HVAC equipment is in the next county. Short of this, a spot should be selected that isolates the inevitable vibration of such equipment from the sound-sensitive area. A good situation is to have the equipment mounted on a concrete pad completely isolated from the structure. In this way, the noise problem is reduced to handling the noise coming through the ducts, a much simpler task than fighting structure-borne vibration.

4.4.2 Identification of HVAC Noise Producers

The various types and paths of HVAC noise producers are identified in Fig. 4-43. This figure provides an interesting study in flanking paths. It is important to remember that there will be relatively little noise reduction unless all of the paths are controlled. A represents the sound room. B represents the room containing the HVAC system. Looking at the noise sources as numbered, 1 and 2 represent the noise produced by the diffusors themselves. The noise is produced by the air turbulence that is created as the air moves through the diffusor. Many diffusors have a noise rating at a given air flow, and the only element of control in this case is selecting the design with the best rating. Don't forget that this applies to the return grille as well as the supply diffusor. Arrows 3 and 4 represent essentially fan noise, which travels to the room via both supply and return ducts and is quite capable of traveling upstream or downstream. The delivery of fan noise over these two paths can be reduced by silencers and/or duct linings. Sizing the ductwork properly is also a means of combating fan noise since sound power output of a fan is fixed largely by air volume and pressure. Arrow 5 represents a good example of a flanking path that is often missed. Depending on how the ceiling in both of the rooms is constructed, the sound from the HVAC unit can travel up through the ceiling in the HVAC room and comes down into room A. Of course the way to control path 5 is to make sure that the ceilings in both rooms are well built, massive enough to control low frequency vibrations, and of course, airtight. Arrow 6 represents that path where the sound can travel through gaps or holes inadvertently left in the partition. This has already been discussed in Section 4.3.11 and in Fig. 4-17. Number 7 represents the sound that can travel straight through a poorly built wall. Numbers 8, 9, and 10 represents the paths that the structure-borne vibrations can take through the structure. We will deal with isolation issues in the next section. Finally, 11 and 12 represent what is called *break-in* noise. This is what happens when sound enters or breaks into a duct and travels down it, radiating into the room.

4.4.3 Vibration Isolation

The general rule is first to do all that can reasonably be done at the source of vibration. The simple act of mounting an HVAC equipment unit on four vibration mounts may help reduce transmitted vibration, may be of no effect at all, or may actually amplify the vibrations, depending on the suitability of the mounts for the job. Of course, if it is successful it would drastically reduce or eliminate paths 8, 9, and 10 in Fig. 4-43. The isolation efficiency is purely a function of the relationship between the frequency of the disturbing source f_d to isolator natural frequency f_n, as shown by Fig. 4-44. If $f_d = f_n$, a resonance condition exists, and maximum vibration is transmitted. Isolation begins to occur when f_d / f_n is equal to or greater than 2. Once in this isolation range, each time f_d / f_n is doubled, the vibration transmission decreases 4–6 dB. It is beyond the scope of this treatment to go further than to identify the heart of both the problem and the solution, leaving the rest to experts in the field.

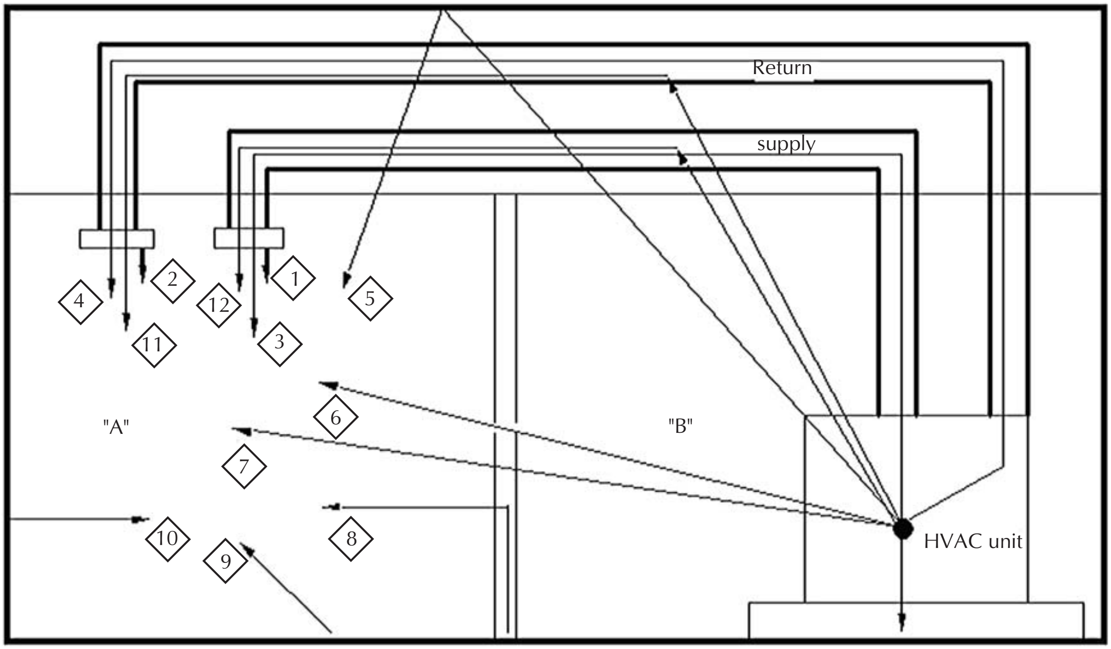


Figure 4-43. Typical paths by which HVAC noise can reach sound-sensitive rooms.

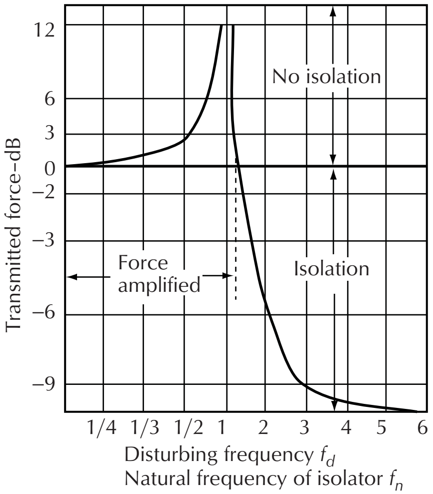


Figure 4-44. Noise from HVAC equipment may be reduced by isolation mounts, or it may actually be amplified. (After ASHRAE, Reference 18.)

4.4.4 Attenuation of Noise in Ducts

Metal ducts with no linings attenuate fan noise to a certain extent. As the duct branches, part of the fan noise energy is guided into each branch. Duct wall vibration absorbs some of the energy, and any discontinuity (such as a bend) reflects some energy back toward the source. A very large discontinuity, such as the outlet of the duct flush with the wall, reflects substantial energy back toward the source. This results in attenuation of noise entering the room, as shown in Fig. 4-45. Unlike many other systems in acoustics this is one attenuation that is greater at low frequencies than at the highs.

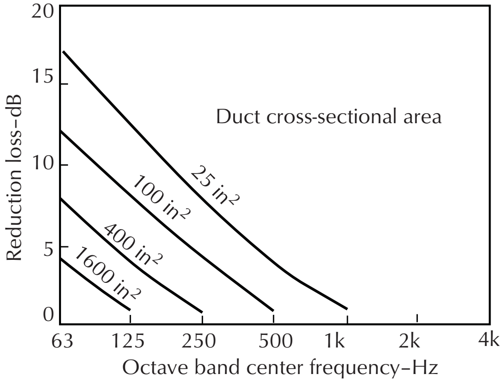


Figure 4-45. The effect of duct cross-sectional area on the attenuation of HVAC noise. (After ASHRAE, Reference 18.)

Lining a duct increases attenuation primarily in the higher audio frequency range. Fig. 4-46 shows measured duct attenuation with 1 inch duct lining on all four sides. The dimensions shown are for the free area inside the duct. This wall effect attenuation is greatest for the smaller ducts. For midband frequencies, a 10 ft length of ducting can account for 40 dB or 50 dB attenuation for ducts 12 inches × 24 inches or smaller. There is a trade-off, however, as decreasing the cross section of the duct increases the velocity of the air moving through it. Higher air velocities produce greater turbulence noise at the grille/diffusor. Great stress is commonly placed on attenuation contributed by right-angle bends that are lined with duct liner. Fig. 4-47 evaluates attenuation of sound in lined bends. Only lining on the sides is effective, which is the way the elbows of Fig. 4-47 are lined. Here again, attenuation is greater at higher audio frequencies. The indicated duct widths are clear measurements inside the lining. The lining thickness is 10% of the width of the duct and extends two duct widths ahead and two duct widths after the bend. It is apparent that the lining contributes much to attenuation of noise coming down the duct, but less so at lower frequencies. Here too, there is a trade-off. Every bend, lined or not, increases the turbulence and therefore the noise.

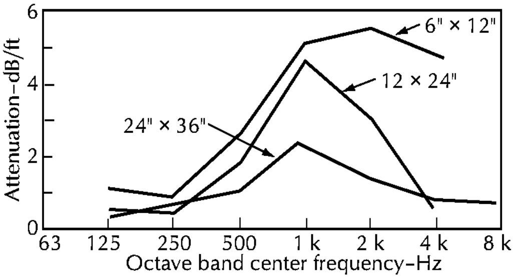


Figure 4-46. Measured noise attenuation in rectangular ducts. (After ASHRAE, Reference 18, which attributes Owens-Corning Fiberglas Corp. Lab Report 32433 and Kodaras Acoustical Laboratories Report KAL-1422-1 submitted to Thermal Insulation Manufacturer's Association.)

4.4.5 Tuned Stub Noise Attenuators

Fan blades can produce line spectra or tonal noise at a blade frequency of

$$\textit{Blade frequency} = \frac{\textit{RPM} \times \textit{Number of blades}}{60\ \text{Hz}} \quad (4\text{-}5)$$

Usually this noise is kept to a minimum when the HVAC engineer selects the right fan. If such tones continue to be a problem, an effective treatment is to install a tuned stub filter someplace along the duct. These can be very effective in reducing fan tones. A typical stub and its attenuation characteristic are shown in Fig. 4-48A. The comparable characteristic of a reactive muffler is also shown in Fig. 4-48B.

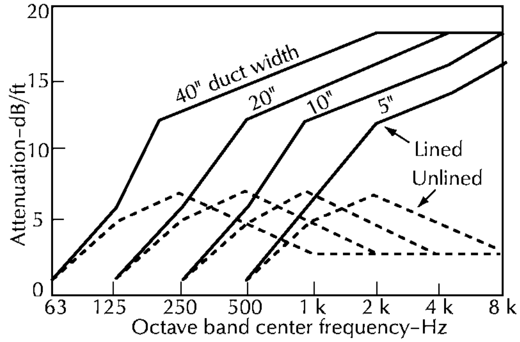
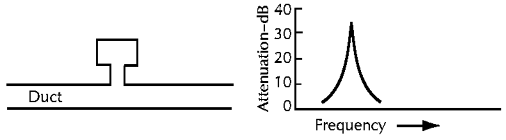
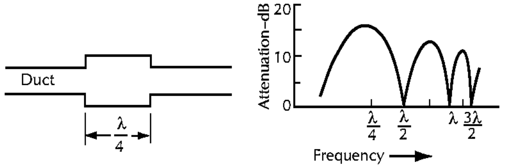


Figure 4-47. Noise attenuation in HVAC square-duct elbows without turning vanes. (After ASHRAE Reference 18.)



A. The tuned stub offers attenuation in a narrow band and is useful in reducing tonal noise from HVAC equipment.



B. The reactive muffler offers a series of attenuation peaks down through the spectrum. These are useful in reduction of specific noise components.

Figure 4-48. The tuned stub and reactive muffler used to attenuate tonal components of noise.

4.4.6 Plenum Noise Attenuators

As previously stated, a most effective procedure in noise reduction is to reduce the noise at, or very close to, the source. If a system produces a noise level that is too high at the sound room end, one possibility is to install a plenum in the supply and another in the return line. Such a plenum is simply a large cavity lined with absorbing material, as shown in Fig. 4-49. Sometimes a nearby room or attic space can be made into a noise-attenuating plenum, usually at the source. The attenuation realized from a plenum can be estimated from the following expression:[19]

$$attenuation = 10\log\left[\frac{1}{S_e\left(\frac{\cos\theta}{2\pi d} + \frac{1-a}{S_w a}\right)}\right] \qquad (4\text{-}6)$$

where,

a is the absorption coefficient of the lining,

S_e is the plenum exit area in square feet,

S_w is the plenum wall area in square feet,

d is the distance between the entrance and exit in feet,

θ is the angle of incidence at the exit (i.e., the angle that the direction d makes with the axis of exit) in degrees.

For those high frequencies where the wavelength is smaller than plenum dimensions, accuracy is within 3 dB. At lower frequencies Eq. 4-5 is conservative, and the actual attenuation can be 5 dB to 10 dB higher than the value it gives.

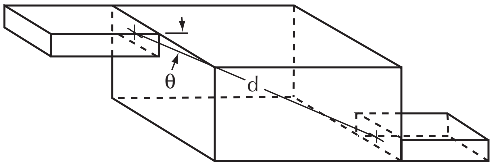


Figure 4-49. A properly designed lined plenum is a very effective attenuator of HVAC noise and is usually located near the equipment. Unused rooms or attic spaces may sometimes be converted to noise-attenuating plenums. (After ASHRAE, Reference 19.)

4.4.7 *Proprietary Silencers*

When space is at a premium and short runs of duct are necessary, proprietary sound-absorbing units can be installed in the ducts at critical points. There are a number of configurations available, and many attenuation characteristics can be expected. The extra cost of such units may be offset by economies their use would bring in other ways. The user should also be aware that silencers produce a small amount of self-noise and care must be taken to allow the air to return to a laminar flow downstream of the silencer.

The general rule is that the air will require a length equal to 10 times the diameter of the duct to regain a laminar flow.

4.4.8 *HVAC Systems Conclusion*

The intent of this HVAC section is to emphasize the importance of adequate attention to the design and installation of the heating, ventilating, and air-conditioning system in the construction of studios, control rooms, and listening rooms. HVAC noises commonly dominate in such sound rooms and are often the focus of great disappointment as a beautiful new room is placed into service. The problem is often associated with the lack of appreciation by the architect and the HVAC contractor of the special demands of sound rooms. It is imperative that an NC clause be written into every mechanical (HVAC) contract for sound-sensitive rooms.

Residential HVAC systems commonly employ small ducts and high velocity air delivery systems. Air turbulence noise increases as the sixth power of the velocity; hence, high velocity HVAC systems can easily be the source of excessive turbulence noise at grilles and diffusers. Keeping air velocity below 400 ft/min for studios and other professional sound rooms is a basic first requirement. Air flow noise is generated at tees, elbows, and dampers; and it takes from 5 to 10 duct diameters for such turbulence to be smoothed out. This suggests that duct fittings should be spaced adequately. Air flow noise inside a duct causes duct walls to vibrate, tending to radiate into the space outside. Thermal duct wrapping (lagging) helps to dampen such vibrations, but even covered, such ducts should not be exposed in sound-sensitive rooms. This oversimplified treatment of HVAC design is meant to underscore the importance of employing expert design and installation talent, not to create instant experts. The overall HVAC project, however, needs the involvement of the audio engineer at each step.[17,18]

References

1. L. L. Beranek, Revised Criteria for Noise in Buildings, *Noise Control*, vol. 3, no. 1, p. 19, January 1957.
2. L. L. Beranek, W. E. Blazier, and J. J. Figwer, Preferred Noise Criteria (PNC) Curves and Their Application to Rooms. Presented at the 81st Meeting of the ASA, Washington DC April 1971.
3. L. L. Beranek, 1989. Balanced Noise Criterion (NCB) curves. *Journal of the Acoustical Society of America* 86: 650 – 664.

4. M. Rettinger, Noise Level Limits in Recording Studios, *dB the Sound Engineering Magazine*, vol. 12, No. 4, pp. 41 – 43, April 1978.
5. B. J. Smith, R. J. Peters, S. Owen, *Acoustics and Noise Control,* Longman Inc. New York 1982.
6. Bruel and Kjaer Instruments Inc. Technical Review # 1-1986.
7. M. Rettinger, *Acoustic Design and Noise Control*, New York: Chemical Publishing Co., Inc., 1973.
8. *Classification for Rating Sound Insulation*, American Society for Testing Materials, designation E413-87.
9. *Classification for Rating Sound Insulation*, ASTM E-413-87 5.4.1/5.4.2.
10. Introduction to Seismic Vibration Control Specifications, Bulletin SVCS-110-3, Mason Industries Inc. 1999.
11. D.W. Green and C. W. Sherry, "Sound Transmission Loss of Gypsum Wallboard Partitions, Report No. 3. 2 X 4 Wood Stud Partitions, *Journal of the Acoustical Society of America*, vol. 71, no. 4, pp. 908 – 914, April 1982. (Report No.1, vol. 71, no. 1 and Report 2, vol. 71, no. 4, also reports extensive measurements of transmission loss.)
12. Architectural Specifications for Isolated Floors, Walls, and Ceilings, Bulletin ACS-100, Hollis, NY: Mason Industries, Inc., 1977.
13. F. A. Everest, *How to Build a Small Budget Recording Studio from Scratch*, Blue Ridge Summit, PA: Tab Books, Inc.
14. F. A. Everest, Glass in the Studio, *dB the Sound Engineering Magazine*, Part I, vol. 18, no. 3, pp. 28 – 33, April 1984: Part II vol. 18, no. 4, pp. 41– 44, May 1984.
15. Breaking The Sound Barrier, Publication AR-3, Libbey-Owens-Ford Co.
16. *ASHRAE Handbook and Product Directory – 1977*. Fundamentals, Chapter 7, "Sound Control Fundamentals," American Society of HVAC Engineers, Inc., 345 East 47th St., New York.
17. *ASHRAE Handbook and Product Directory – 1976*. Systems, Chapter 35, "Sound and Vibration Control," American Society of HVAC Engineers, Inc., 345 East 47th St., New York.
18. *ASHRAE Handbook – 1985*. Fundamentals, Section 2.0, Chapter 7, "Sound and Vibration Fundamentals," American Society of HVAC Engineers, Inc., 345 East 47th St., New York.

Chapter 5

Small Room Acoustics

by Jeff Szymanski

5.1 Acoustical Treatment Overview 95
5.2 Acoustical Absorption 95
  5.2.1 Absorption Testing 96
    5.2.1.1 Reverberation Chamber Method 96
    5.2.1.2 Impedance Tube Testing Methods 98
    5.2.1.3 Other Absorption Testing Methods 98
    5.2.1.4 Absorption Ratings 99
    5.2.1.5 Interpreting Test Results 100
  5.2.2 Porous Absorbers 100
    5.2.2.1 Mineral and Natural Fibers 101
    5.2.2.2 Acoustical Foams 103
    5.2.2.3 Acoustical Tiles 105
    5.2.2.4 Spray and Trowel Applied Treatments 105
    5.2.2.5 Carpet and Draperies 105
  5.2.3 Discrete Absorbers 106
    5.2.3.1 People and Seats 106
    5.2.3.2 Acoustical Baffles and Banners 106
    5.2.3.3 Other Furnishings and Objects 107
  5.2.4 Resonant Absorbers. 107
    5.2.4.1 Membrane Absorbers 107
    5.2.4.2 Helmholtz Resonators 109
    5.2.4.3 Perforated Membrane Absorbers 109
    5.2.4.4 Slat Absorbers 110
  5.2.5 Bass Traps 111
  5.2.6 Applications of Absorption 112
    5.2.6.1 Use of Absorption in Reverberant Spaces 112
    5.2.6.2 Use of Absorption in Nonreverberant Spaces 112
  5.2.7 Subjective Impact of Absorption. 113
  5.2.8 Absorption and Absorption Coefficients of Common Building Materials and Absorbers 113
5.3 Acoustical Diffusion 113
  5.3.1 Diffuser Testing: Diffusion, Scattering, and Coefficients. 115
  5.3.2 Mathematical (Numerical) Diffusers. 116
  5.3.3 Random Diffusion 117
  5.3.4 Applications of Diffusion 118
5.4 Reflection and Other Forms of Sound Redirection 118
5.5 Electronic Treatments 119
5.6 Acoustical Treatments and Life Safety 119
5.7 Acoustical Treatments and the Environment 120
References 121

Bibliography 122

5.1 Acoustical Treatment Overview

It is possible that there is no area in professional audio where there is more confusion, folklore, and just plain misinformation than in the area of acoustical treatment. Everyone, it seems, is an acoustical expert. Of course, like most disciplines, much of acoustics is logical and intuitive if one understands the fundamentals. As Don Davis wrote, "In audio and acoustics the fundamentals are not difficult; the physics are."[1] The most fundamental of all rules in acoustics is that nothing is large or small. Everything is large or small relative to the wavelength of the sound under consideration. This is one of the realities that makes the greater field of audio so fascinating. Human ears respond to a range of wavelengths covering approximately 10 octaves, as compared to eyes, which respond to a range of frequencies spanning about one octave. Even though the bandwidth of visible light is obviously much larger than that of audible sound because of the much higher frequencies involved, the range of the wavelengths in this 10 octave bandwidth poses some unique challenges to the acoustician. We must be able to deal with sounds whose wavelengths are 17 m (56 ft) and sounds whose wavelengths are 1.7 cm (0.6 in).

Getting rooms to sound good is an art as much as it is a science. In some situations, concert halls, for example, there is a reasonable agreement as to what makes for a good hall. In other applications, such as home theaters, recording studios, or houses of worship, there is little agreement among the users, let alone the consultants, as to how these rooms should sound. Considerable research must be done before we are able to trace all of the subjective aspects of room acoustics back to physical parameters. However, some fundamental rules and principles can be noted. The acoustician has very few tools. In fact, there are only two things one can do to sound. It can either be absorbed or redirected, Fig. 5-1. Every room treatment, from a humble personal listening room to the most elaborate concert hall, is made up of materials that either absorb or redirect sound. Room acoustics boils down to the management of reflections. In some situations, reflections are problems that must be removed. In other situations, reflections are purposely created to enhance the experience.

This chapter will address general issues of modifying the way rooms sound. Absorption and absorbers will be covered in detail, as well as diffusion and diffusers, and other forms of sound redirection. Additionally, some discussion on the controversial topic of electroacoustical treatments, and brief sections that touch on life safety and the environment as they pertain to acoustical treatments are provided. The information will be thorough, but not exhaustive. There are, after all, entire books dedicated to the subject of acoustical treatments.[2] The intention here is to be able to provide a solid understanding of the fundamentals involved. Specific applications will be dealt with in subsequent chapters.

5.2 Acoustical Absorption

Absorption is the act of turning acoustical energy into some other form of energy, usually heat. The unit of acoustical absorption is the sabin, named after W.C. Sabine (1868–1919), the man considered the father of modern architectural acoustics. It is beyond the scope of this treatment to tell the story of Sabine's early work on room acoustics, but it should be required reading for any serious student of acoustics. Theoretically, 1.0 sabin equates to one square meter (m^2) of complete absorption. Sabine's original work involved determining the sound absorbing power of a material. He posited that comparing the performance of a certain area of material to the same area of open window would yield its absorbing power relative to the ideal.[3] For example, if 1.0 m^2 of a material yielded the same absorbing power as 0.4 m^2 of open window, the relative absorbing power—what we now call the absorption coefficient—would be equal to 0.4.[4]

How absorption is used depends on the application and the desired outcome. Most of the time, absorption is used to make rooms feel less live or reverberant. Absorber performance varies with frequency, with most working well only over a relatively narrow range of frequencies. In addition, absorber performance is not necessarily linear over the effective frequency range.

Measuring or classifying absorbers is not as straightforward as it may seem. There are two main laboratory methods: the impedance tube method and the reverberation chamber method, both of which will be discussed in detail below. Field measurement of absorption will also be discussed below. Absorber performance can also be determined theoretically; discussions of those methods are beyond the scope of this chapter. (The reader is referred to the Bibliography at the end of this chapter for advanced absorber theory texts.)

There are three broad classifications of absorbers: porous, discrete, and resonant. While it is not uncommon for people to design and build their own absorbers (indeed, there has been something of a resurgence in do-it-yourself absorber construction in recent years as a result of the proliferation of how-to guides and Internet discussion forums—this information may or may not be reliable, depending on the reliability of the online resource and the relative expertise of the "experts"

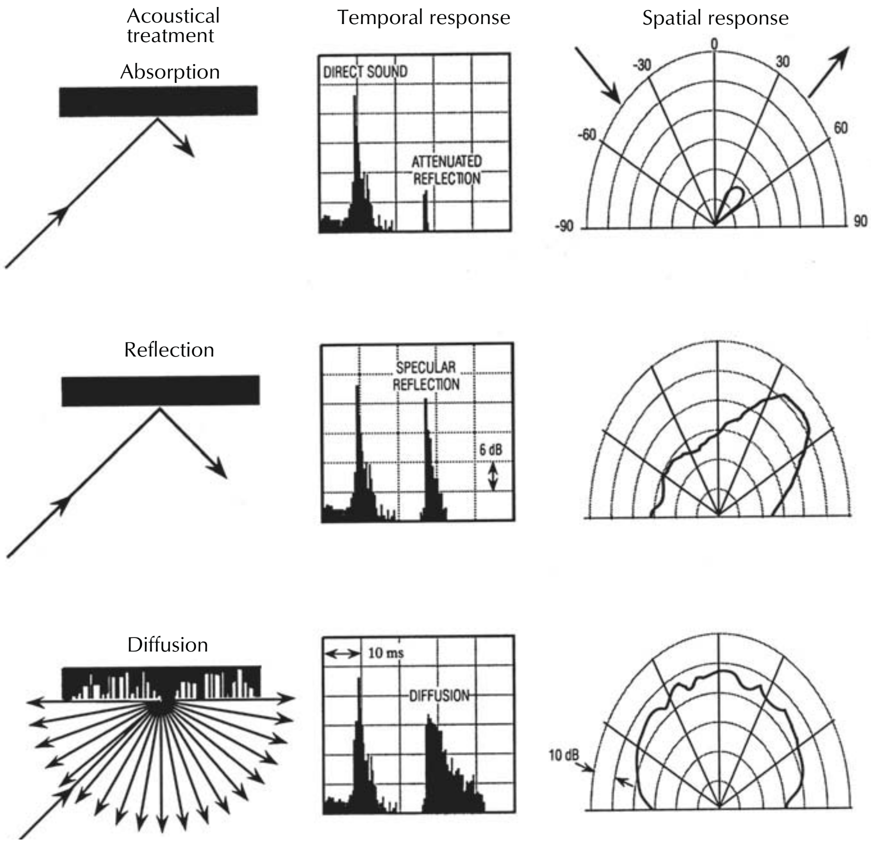


Figure 5-1. Comparison of absorption reflection and diffusion.

offering guidance), many excellent porous and resonant absorbers are available commercially. Fundamental information about the design of absorbers is included here for two reasons: there may be those who want to build their own absorbers, and more importantly, these absorbers are sometimes inadvertently constructed in the process of building rooms. This is especially true of resonant absorbers.

5.2.1 Absorption Testing

Standardized testing of absorption began with Sabine and continues to be developed and improved upon in the present day. As mentioned above, the two standardized methods for measuring absorption are the reverberation chamber method and the impedance tube method. One can also measure absorption in the field by using either the standardized methods or the other techniques discussed below.

5.2.1.1 Reverberation Chamber Method

The work of Sabine during the late 19th and early 20th centuries is echoed in the present-day standard methods for measuring absorption in a reverberation chamber: ASTM C423 and ISO 354.[5,6] In both methods, the general technique involves placing a sample of the material to be tested in a reverberation chamber. This is a chamber that has no absorption whatsoever. The rate of sound decay of the room is measured with the sample in place and compared to the rate of sound decay of the empty room. The absorption of the sample is then calculated.

The method of mounting the sample in the test chamber has an effect on the resulting absorption. Thus, standardized methods for mounting are provided.[6,7] The most common mounting methods employed are Types A, B, and E. Type A simply involves placing the test sample—usually a board-type wall or ceiling absorber—flat against the predefined test area in the chamber (typically on the floor). Type B mounting is typically encountered with acoustical materials that are spray or trowel applied. The material is first applied to a solid backing board and then tested by placing the treated boards over the predefined test area in the chamber. Type E mounting is the standard method employed for absorbers such as acoustical ceiling tiles. This mounting includes a sealed air space of a defined depth behind the absorbers to mimic the real-world installation of acoustical ceiling tiles with an air plenum above. The depth is defined in millimeters and is denoted as a suffix. For example, a test of acoustical ceiling tiles in an E400 mounting means that the tiles were tested over sealed air space that was 400 mm (16 in) deep.

It should be noted that Type A mounting for board-type wall and ceiling absorbers is so often used as the default method that any mention of mounting method is often carelessly omitted in manufacturer literature. Regardless, it is important to verify the mounting method used when evaluating acoustical performance data. If there is any uncertainty, a complete, independent laboratory test report should be requested and evaluated. Details of the mounting method must be included in the lab report to fulfill the requirements of the test standards.

ASTM C423 is generally used in certified North American laboratories; ISO 354 is generally the adopted standard in European countries. The methods are very similar, but there are some noteworthy differences that can yield different testing results. A main difference that is a frequent subject of criticism is the different minimum sample sizes. The minimum area of material when testing board-type materials in accordance with ASTM C423 is 5.6 m^2 (60 ft^2)[5] (the recommended test area is 6.7 m^2 [72 ft^2]) and that of ISO 354 is 10 m^2 (107.6 ft^2)[6]. In general, this difference in sample size can result in a material having slightly lower absorption coefficients when tested in accordance with ISO 354 relative to the same material tested in accordance with ASTM C423. The ISO method is generally regarded as a more realistic approach when the test results are being applied to spaces that are larger than the test chamber, as is often the case. Nonetheless, ASTM test results have been widely and successfully used in architectural acoustic room design applications for many decades.

The reverberation chamber methods can also be applied to discrete absorbers, such as auditorium seating, highway barriers, office partitions, and even people. The main difference between testing the discrete absorbers and testing panel-type absorbers is how the results are reported. If a material occupies a commensurable area of a test chamber surface, absorption coefficients can be calculated. By contrast, the results of a test of some number of discrete absorbers are generally reported in sabins/unit. (Sometimes referred to as *Type J mounting* in the literature, provided the test met the standard requirements for that mounting). For example, the absorption of acoustical baffles—the type that might be hung from a factory or gymnasium ceiling—is typically reported in sabins/baffle.

When calculating absorption coefficients for board-type absorbers, the number of sabins in each frequency band is divided by the surface area of the test chamber covered by the sample material. The resulting quantity is the Sabine absorption coefficient, abbreviated α_{SAB}. The vast majority of absorption coefficients reported in the literature is Sabine absorption coefficients. Since the material is tested in a reverberant space, the Sabine absorption coefficients are useful for predetermining the acoustical properties of a space, provided that the product is intended for use in a similarly reverberant space (i.e., a space where sound can be considered to be impinging equally on a surface from all angles of incidence).

The frequency range of reverberation chamber measurements is limited. At low frequencies, modal effects can dominate the test chamber, thus making accurate measurements of sound decay difficult. At high frequencies, the chambers are large enough that the absorption of air will start to affect the measurement results. Therefore, the frequency range for a reverberation chamber test is typically limited to the ⅓ octave bands between 100 and 5000 Hz. This is sufficient for most materials and applications as it spans a full six octaves over what is commonly referred to as the *speech* range of frequencies—i.e., the range of frequencies that are important to address design issues related to speech communication.

When acoustical treatments are specifically designed to absorb low frequencies, the reverberation chamber method can fall short. However, D'Antonio has implemented a special application of the ASTM C423 method that utilizes fixed microphone positions (vs, the more typical rotating microphone) that measure the decay of the actual modal frequencies of the room. Using this method, D'Antonio has been able to measure low frequency absorption down to the 63 Hz octave band.[8,9] The impedance tube method (discussed below) can also be used to measure low frequency absorption, but a large tube with heavy walls (such as poured concrete) is required.

5.2.1.2 Impedance Tube Testing Methods

The laboratory methods generally involve the use of an impedance tube to measure absorption of normally incident sound—i.e., sound arriving perpendicular to the sample. There are two standard methods to measure absorption in an impedance tube: the single-microphone, standing wave method; and a two-microphone, transfer function method.[2] In general, impedance tube measurements are relatively inexpensive, relatively simple to perform, and can be very useful in research and development of absorber performance. In the standing wave method, for example, the normal absorption coefficient (α_n) can be calculated from

$$\alpha_n = \frac{I_i - I_r}{I_i} \tag{5-1}$$

where,

I_i is the incident sound intensity,

I_r is the reflected sound intensity.

While the cost and time saving benefits of the impedance tube method are obvious, care should be taken since the normal absorption coefficients are not equivalent to the Sabine absorption coefficients discussed in the previous section. In fact, unlike α_{SAB}, α_n can never be greater than 1.0. In one set of experiments, α_{SAB} was as little as 1.2 times and as much as almost 5.0 times greater than α_n.[10] Regardless, there is no established empirical relationship between α_{SAB} and α_n. Normal absorption coefficients should not be used to calculate the properties of a space using standard reverberation time equations.

One main advantage offered by normal absorption coefficients is that they offer an easy way to compare the performance of two absorbers. Reverberation chambers have inherent reproducibility issues (explained in more detail below). The impedance tube can overcome this to some extent. One limitation of the impedance tube is frequency range; large tubes are needed to test low frequencies. Another is that tests of resonant absorbers tend not to produce accurate results, because of the small sample size.

5.2.1.3 Other Absorption Testing Methods

Many methods can be employed for the measurement of sound absorption outside the confines of a laboratory test chamber or impedance tube.[2] Of course, both the reverberation chamber method and the impedance tube methods can be adopted for use in the field. In fact, Appendix X2 of ASTM C423 provides guidelines for carrying out the reverberation method in the field.[5]

When the sound impinging on an absorber is not totally random—as is the case, more often than not—there may be better methods for describing its performance. One of these methods, described by Brad Nelson,[11] involves the analysis of a single reflection by means of signal processing techniques. Although Nelson's method describes the measurement of absorption at normal incidence, his method can be extended to determine the in situ absorption coefficients of a material at various angles of incidence, which can be particularly useful for the analysis of absorbers that are being used for reflective control in small rooms. Nelson's method was employed by the author to determine the in situ angular absorption coefficient (α_θ) of two different porous absorbers, the results of which are shown graphically in Fig. 5-2 for reflections in the 2000 Hz band. The results at least partly confirm what has often been observed in recording studios: sculpted acoustical foam tends to be more consistent in its control of reflections at oblique angles of incidence relative to flat, fabric-covered, glass fiber panels of higher density. Or, to put it another way, the glass fiber panel offers more off-axis reflections than the acoustical foam panel. Of course, the relative merits of one acoustical treatment over the other are subjective. The important point is that the differences are quantifiable.

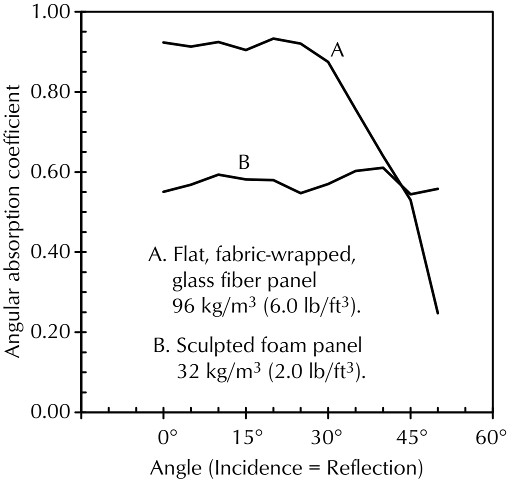


Figure 5-2. Angular absorption coefficients (α_θ) of two absorbers for the 2000 Hz $^1/_3$-octave band.

5.2.1.4 Absorption Ratings

There are three single number ratings associated with absorption, all of which are calculated using the Sabine absorption coefficients. The first and most common is the *Noise Reduction Coefficient* (NRC). The NRC is the arithmetic average of the 250, 500, 1000, and 2000 Hz octave-band Sabine absorption coefficients, rounded to the nearest 0.05.[5] The NRC was originally intended to be a single number rating that gave some indication of the performance of a material in the frequency bands most critical to speech.

To partly address some of the limitations of the NRC, the *Sound Absorption Average* (SAA) was developed.[5] Similar to NRC, the SAA is an arithmetic average, but instead of being limited to four octave bands, the Sabine absorption coefficients of the twelve ⅓ octave bands from 200 through 2500 Hz are averaged and rounded to the nearest 0.01. Table 5-1 provides an example calculation of both NRC and SAA for a set of absorption coefficients.

Table 5-1. Sample Sabine Absorption Coefficient (α_{SAB}) Spectrum with Corresponding Single Number Ratings, NRC, SAA, and α_w.

⅓ Octave Band Center Frequency	α_{SAB}	⅓ Octave Band Center Frequency	α_{SAB}
100 Hz	0.54	1250 Hz	0.39
125 Hz	1.38	1600 Hz	0.31
160 Hz	1.18	2000 Hz	0.30
200 Hz	0.88	2500 Hz	0.23
250 Hz	0.80	3150 Hz	0.22
315 Hz	0.69	4000 Hz	0.22
400 Hz	0.73	5000 Hz	0.20
500 Hz	0.56		
630 Hz	0.56	**NRC =**	**0.55**
800 Hz	0.51	**SAA =**	**0.53**
1000 Hz	0.47	α_w **=**	**0.30 (LM)**

Finally, ISO 11654 provides a single number rating for materials tested in accordance with ISO 354 called the weighted sound absorption coefficient (α_w).[12] A curve matching process is involved to derive the α_w of a material. Additionally, shape indicators can be included in parentheses following the α_w value to indicate areas where absorption has significantly exceeded the reference curve. Table 5-1 shows the α_w for the set of absorption coefficients, with the LM indicating that there may be excess low and mid-frequency absorption offered that is not otherwise apparent from the α_w value. This is useful in that it indicates the actual octave-band or ⅓ octave band absorption coefficients are probably worth looking into in greater detail.

None of the metrics described above gives an accurate representation of the absorptive behavior (or lack thereof) of a material. NRC averages four bands in the speech frequency range. The problem, of course, is that many different combinations of four numbers can result in the same average, as shown in Table 5-2. The same can be said for SAA. Nonetheless, NRC and SAA can be compared to give a little bit more information than each rating gives on an individual basis. If NRC and SAA are very close, the material probably does not have any extreme deviations in absorption across the speech range of frequencies. If SAA is drastically different from NRC, it may be indicative of some large variations at certain ⅓ octave bands. These are, of course, only single number ratings; none of them takes into account the performance of the material below the 200 Hz ⅓ octave band. They can, at most, provide a cursory indication of the relative performance of a material. A full evaluation of the performance of a material should always involve looking at the octave or ⅓ octave band data in as much detail as possible.

Table 5-2. Two Different Sample Sabine Absorption Coefficient (α_{SAB}) Spectra with Equal NRC and SAA.

	α_{SAB}	
⅓ Octave Band Center Frequency	Material 1	Material 2
100 Hz	0.54	0.01
125 Hz	1.38	0.01
160 Hz	1.18	0.09
200 Hz	0.88	0.18
250 Hz	0.80	0.33
315 Hz	0.69	0.39
400 Hz	0.73	0.42
500 Hz	0.56	0.57
630 Hz	0.56	0.58
800 Hz	0.51	0.67
1000 Hz	0.47	0.73
1250 Hz	0.39	0.69
1600 Hz	0.31	0.60
2000 Hz	0.30	0.58
2500 Hz	0.23	0.65
3150 Hz	0.22	0.67
4000 Hz	0.22	0.80
5000 Hz	0.20	0.77
NRC =	**0.55**	**0.55**
SAA =	**0.53**	**0.53**

5.2.1.5 Interpreting Test Results

As mentioned at the beginning of this chapter, the acoustical treatment industry is rife with misinformation. Test results, sadly, are no exception. Information in manufacturer literature or on their Web sites is fine for evaluating materials on a cursory basis. This information should eventually be verified, preferably with an independent laboratory test report. If manufacturers cannot supply test reports, any absorption data reported in their literature or on their Web sites should be treated as suspect.

When absorption data is evaluated, the source of the data should be understood, both in terms of which standard method was used and which independent test laboratory was used. Again, the test reports can help clear up any confusion. Close attention should be paid to subtle variations in test results, such as a manufacturer who tested the standard-minimum area of material in lieu of the standard-recommended area of material for an ASTM C423 test. If two materials are otherwise similar, a variation in sample size could explain some of the variation in measured absorption.

Additionally, there are reproducibility issues with the reverberation chamber method. Saha has reported that the absorption coefficients measured in different laboratories vary widely, even when all other factors—e.g., personnel, material sample, test equipment, etc.—are kept constant.[13] Cox and D'Antonio have found absorption coefficient variations between laboratories to be as high as 0.40.[2]

Finally, it is worth noting that Sabine absorption coefficients will often exceed 1.00. This is a source of great confusion since theory states that absorption can only vary between 0.00 (complete reflection) and 1.00 (complete absorption). However, the 0 to 1 rule only applies to, for example, normal absorption coefficients, which are calculated using the measurement of direct versus reflected sound intensity. Sabine absorption coefficients, remember, are calculated using differences in decay rate and by dividing the measured absorption by the sample area. In theory, this should still keep the Sabine absorption coefficients below 1.0. However, edge and diffraction effects are present and are frequently cited (along with some nominal hand-waving) to explain away values greater than 1.0. Edge and diffraction effects are true and valid explanations,[14] but can be confusing in their own right. For example, samples are often tested with the edges covered—i.e., not exposed to sound. Absorption coefficients greater than 1.0 resulting from such a test can therefore be attributed mainly to diffraction effects, which is the process where sound that would not normally be incident on a sample is bent towards the sample and absorbed. The confusion arises when these test results are utilized in applications where the edges of the sample will be exposed to sound.

A better explanation might be simply that Sabine absorption coefficients are not percentages. The variables in the calculation of the Sabine absorption coefficient are rate of decay and test sample area. A change in the former divided by the latter is basically what is being determined, which does not strictly conform to the definition of a percentage. Based on this explanation, an α_{SAB} value greater than 1.0 simply indicates a higher absorption than a value lower than 1.0, all other factors being equal. For example, a material with a Sabine absorption coefficient of 1.05 at 500 Hz will absorb more sound at 500 Hz than the same area of a material having a Sabine absorption coefficient of 0.90, provided that both materials were tested in the same manner.

Regardless of the validity of Sabine absorption coefficients greater than 1.0, they are usually rounded down to 0.99 for the purposes of predictive calculations. This rounding down is especially important if, for example, equations other than the Sabine equation are used to determine reverberation time. Of course, there has been ample debate about this rounding. For example, technically it is not rounding but scaling that is being done. As Saha has pointed out, why only scale the numbers greater than 1.0—what's to be done, if anything, with the other values?[13]

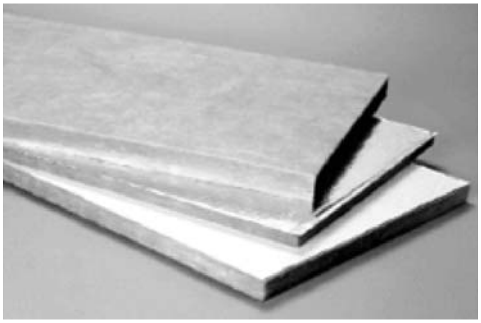
5.2.2 Porous Absorbers

Porous absorbers are the most familiar and commonly available kind. They include natural fibers (e.g., cotton and wood), mineral fibers (e.g., glass fiber and mineral wool), foams, fabrics, carpets, soft plasters, acoustical tile, and so on. The sound wave causes the air particles to vibrate down in the depths of porous materials, and frictional losses convert some of the sound energy to heat energy. The amount of loss is a function of the density or how tightly packed the fibers are. If the fibers are loosely packed, there is little frictional loss. If the fibers are compressed into a dense board, there is little penetration and more reflection from the surface, resulting in less absorption.

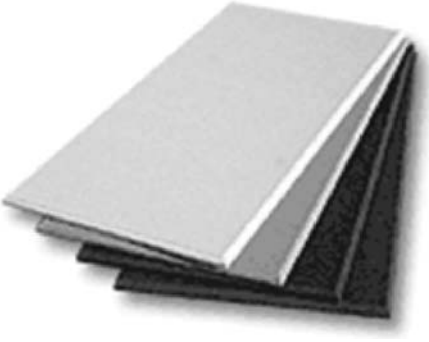
Mainly because there is a veritable plethora of extant information with which to work, the Owens Corning 700 Series of semi-rigid glass fiber boards will be discussed in the next section to not only highlight one of the more popular choices for porous absorber, but also to illustrate various trends—such as absorption dependence on thickness and density—that are not uncommon with porous absorbers in general.

5.2.2.1 Mineral and Natural Fibers

Of the varieties of mineral fiber, one of the most popular is the glass fiber panel or board, Fig. 5-3. The absorption of sound for various densities of Owens Corning 700 Series boards is shown in Figs. 5-4 and 5-5.[15] Fig. 5-4 shows the absorption for 2.5 cm (1 in) thick boards. None of the three densities absorbs well at frequencies below 500 Hz. At the higher audio frequencies, the boards of 48 kg/m^3 and 96 kg/m^3 (3.0 lb/ft^3 and 6.0 lb/ft^3, respectively) densities are slightly better than the lower density 24 kg/m^3 (1.5 lb/ft^3) material. Fig. 5-5 shows a comparison between different densities of the 10.2 cm (4.0 in) thick fiberglass boards. In this case, there is little difference in absorption between the three densities.[15]



A. Raw material.



B. Fabric finished panels.

Figure 5-3. Glass fiber absorbers.

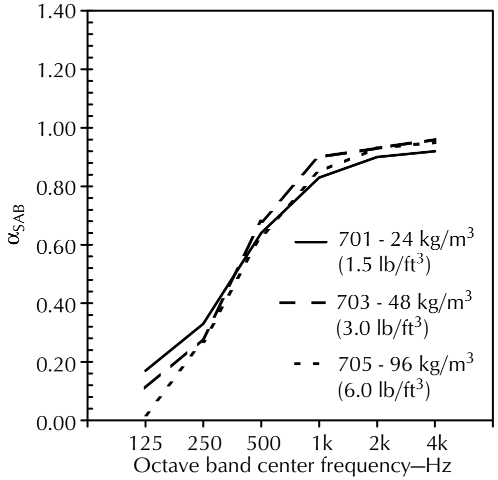


Figure 5-4. The effect of density on the absorption of Owens Corning 700 series glass fiber boards of 2.5 cm (1 in) thickness, Type A mounting.[15]

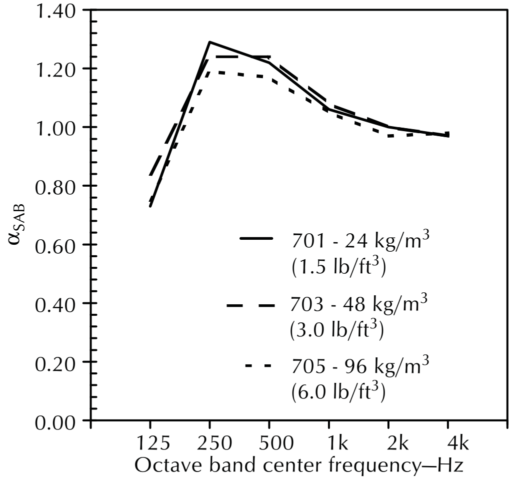


Figure 5-5. The effect of density on the absorption of Owens Corning 700 series glass fiber boards of 10.2 cm (4 in) thickness.[15]

Boards of medium density have a mechanical advantage in that they can be cut with a knife and press-fitted into place. This is more difficult with materials that have a 24 kg/m^3 (1.5 lb/ft^3) density and lower, such as building insulation. The denser the board, the greater the cost. Most acoustical purposes are well served by glass fiber of 48 kg/m^3 (3.0 lb/ft^3) density, although some consultants specify a 96 kg/m^3 (6.0 lb/ft^3) material. A number of consultants regularly specify absorbers that are composed of multiple densities, for example, a combination of Owens Corning 701, 703, and 705. In theory, a multidensity absorber (assuming the least dense material is exposed to the sound source with gradually increasing densities toward the wall) will be as good as or better than a single-density absorber of the same thickness.[16] In practice, this tends to hold true.

Fig. 5-6 explores the effect of thickness of 703 Fiberglas on absorption. The absorption of low-frequency sound energy is much greater with the thicker maerials.[15]

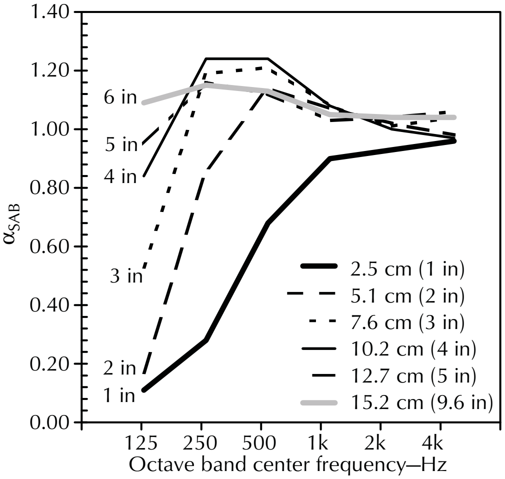


Figure 5-6. The effect of thickness on the absorption of Owens Corning 703 glass fiber boards 48 kg/m^3 (3 lb/ft^3), Type A mounting.[15]

Fig. 5-7 shows the effect of air space behind a 2.5 cm (1 in) thick Owens Corning Linear Glass Board. As the air space is increased in steps from 0 to 12.7 cm (0 to 5 in), the lower-frequency absorption increases progressively.[15] It is sometimes cost-effective to use thinner glass fiber and arrange for air space behind it; it is sometimes cost-effective to use glass fiber of greater thickness. At other times, the need for low-frequency absorption is so great that both thick material and air space are required.

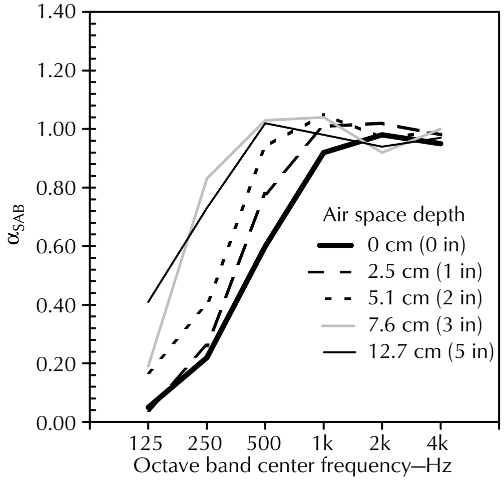
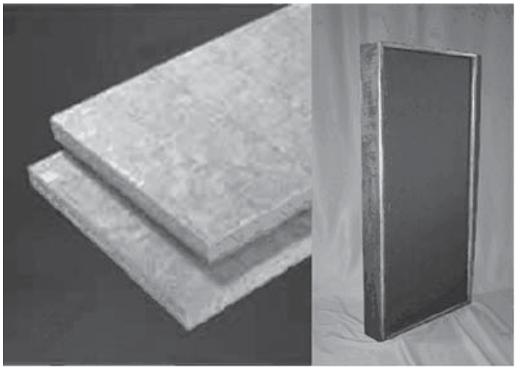


Figure 5-7. The effect of mounting over an air space on the absorption of Owens Corning Linear Glass Cloth faced board of 2.5 cm (1 in) thickness.[15]

In acoustical applications, mineral wool (or rock wool) is another popular variation of mineral fiber board, Fig. 5-8. Figs. 5-9 and 5-10 provide an overview of absorption coefficients for materials available from Roxul.[17] The main difference between glass fiber and mineral wool is that mineral wool is generally made from basalt (glass fiber comes from silicates), which leads to a higher heat tolerance.



A. Raw material. B. Framed and fabric-finished panel.

Figure 5-8. Mineral wool absorbers.

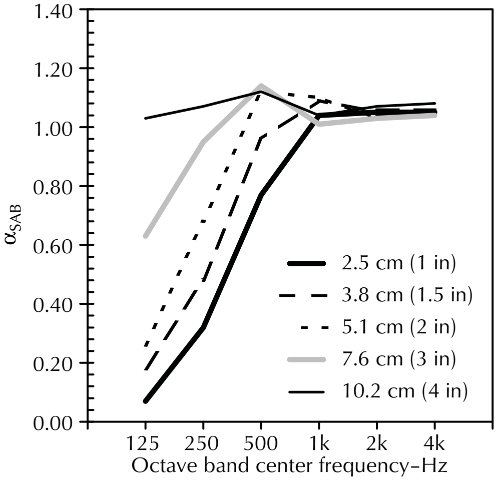


Figure 5-9. The effect of thickness on the absorption of Roxul RockBoard 40 64 kg/m^3 (4 lb/ft^3) mineral wool boards.[17]

Natural fiber materials used in acoustical applications include wood fibers and cotton fibers. Tectum, Inc. manufactures a variety of ceiling and wall panels from aspen wood fibers, which produces a durable acoustical treatment. Absorption coefficients for some Tectum, Inc. materials are shown in Fig. 5-11.[18] There are also an

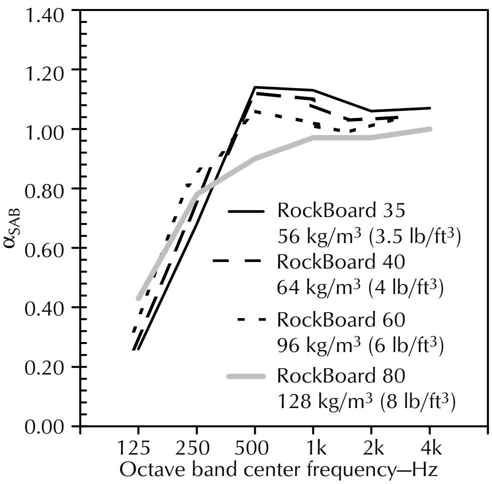


Figure 5-10. The effect of density on the absorption of Roxul RockBoard mineral wool boards of 5.1 cm (2 in) thickness.[17]

increasing number of suppliers of natural cotton absorption panels. The absorption of natural cotton panels—so far as they have been developed—appears to be comparable to mineral fiber panels of similar density.

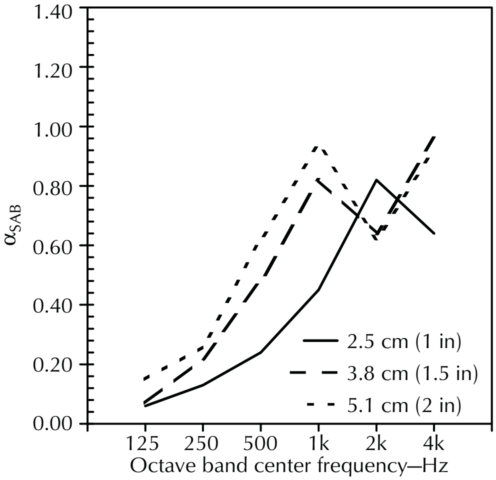


Figure 5-11. The absorption of different thicknesses of Tectum Wall Panels, Type A mounting.[18]

Most fibrous absorbers will be covered with some sort of acoustically transparent fabric finish that is both decorative and practical. The fabric finish is decorative because the natural yellow or green of the glass fiber and mineral wool panels tends to be less than aesthetically pleasing; the finish is practical because airborne fibers from mineral fiber materials can be breathing irritants. Perforated metal (with or without powder-coated finish) and plastic coverings with a high percent of open area (much higher than resonant perforated absorbers discussed below) can also be used with fibrous absorbers. Perforated coverings are typically employed for decorative purposes, maintenance purposes, or to protect the panels from high impacts, such as might occur in a gymnasium. Foil and paper finishes are also sometimes available as low-cost means of containing fibers for glass fiber or mineral wool panels. Because of reflections from the foil or paper, the high-frequency absorption of the faced side of the absorber is significantly lower than that of the unfaced side. (The thin foil or paper used is sometimes referred to as a *membrane*. This has led to confusion with resonant membrane, or diaphragmatic absorbers. For clarity, foil or paper facings as they are applied to fibrous absorber panels are not resonant membranes in the strict sense, but do provide some nominal increases in low-frequency absorption when the foil or paper is exposed to the incident sound.)

To provide some impact resistance, as well as to provide a surface conducive for some office applications (such as for office partitions), a thin (usually 3 mm) glass fiber board of high density (usually 160 to 290 kg/m³ [10 to 18 lb/ft³]) can be applied over the face of a fibrous absorber before the fabric finish is applied. This is often referred to as a tackable surface finish since it can readily accept push pins and thumbtacks.

In terms of installation ease, natural fibers hold some promise since they will offer relief from the itch associated with the handling of mineral fiber boards. Natural fiber products can also be installed without covering, and Tectum, Inc. states that their wood fiber panels can be repainted several times without significant degradation of acoustical performance.

5.2.2.2 Acoustical Foams

There are various types of reticulated open cell foams for acoustical applications, Fig. 5-12. Closed cell foams also find applications in acoustics, but largely as substrates from which acoustical diffusers can be formed. The most common foams used as open cell acoustical absorbers in architectural applications are polyurethane (esters and ethers) and melamine foams. Unlike fibrous boards, foam panels are easy to cut and can be sculpted into shapes and patterns. Besides the ubiquitous wedges and pyramids, acoustical foams have been created with various square, saw tooth, and even curved patterns sculpted into the faces. While removing material generally serves to decrease absorption, creating more exposed surface area tends to increase it. Figs. 5-13 and 5-14 provide absorption coefficients for different patterns of foam and different thicknesses of foam of the same pattern, respectively, for acoustical foam panels available from Auralex Acoustics, Inc.[19]

Figure 5-12. Open cell polyurethane acoustical foam.

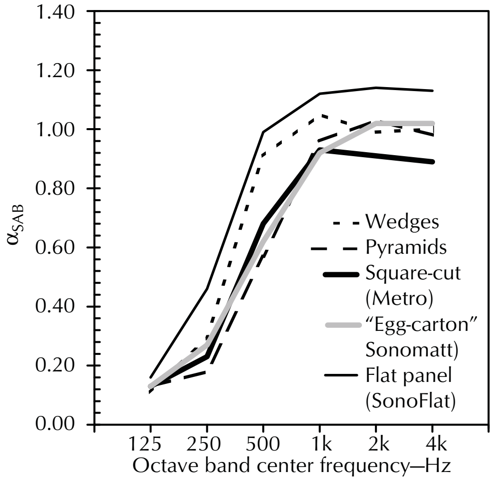


Figure 5-13. The effect of shape on the absorption of Auralex Acoustics polyurethane foam panels of 5.1 cm (2 in) thickness, Type A mounting.[19]

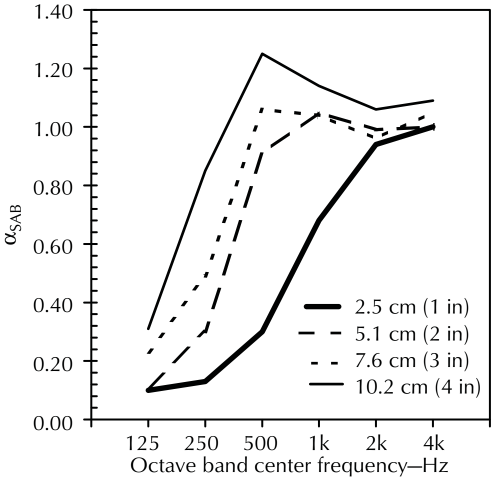


Figure 5-14. The effect of thickness on the absorption of Auralex Acoustics Studiofoam Wedges, Type A mounting.[19]

In general, acoustical foams are of lower density than fibrous materials; acoustical foam densities are generally in the 8.0 to 40 kg/m³ (0.5 to 2.5 lb/ft³) range. This means that mineral fiber panels tend to provide higher absorption coefficients than foam panels of the same thickness. However, acoustical foams can generally be installed without any decorative covering, which can make them more cost-effective—mineral fiber panels tend to require a fabric finish, or some other cover to contain airborne fibers. Melamine foams, such as the acoustical products offered by Pinta Acoustic, Inc. (formerly Illbruck) are white in color and have a higher resistance to fire relative to polyurethane foams. However, melamine foams generally have lower absorption coefficients (largely due to lower densities) and tend to be less flexible, making them more prone to damage than polyurethane foams. A sampling of the acoustical performance of some melamine foam products available from Pinta Acoustic, Inc. is provided in Fig. 5-15.[20] Melamine foams may be painted (the manufacturer should always be consulted about this), while polyurethane foams should generally not be painted. Because of this, companies offering polyurethane foams generally have a wider variety of colors available.

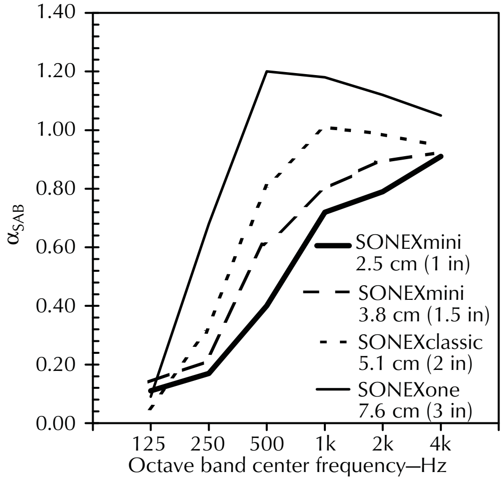


Figure 5-15. The absorption of Pinta Acoustic melamine foam panels of different thicknesses, Type A mountings.[20]

5.2.2.3 Acoustical Tiles

Acoustical tiles have the highest density of the porous absorbers. They are widely used for suspended (lay-in) ceiling treatments. Years ago, it was common to see 30 cm × 30 cm (12 in × 12 in) tile mounted directly to a hard plaster (Type A mounting). This is not a very efficient way to use this type of absorber and is no longer popular.

The standard sizes for acoustical tiles are 61 cm square (24 in × 24 in) or 61 cm × 122 cm (24 in × 48 in) and the Sabine absorption coefficients are usually given for Type E400 mounting, which mimics a lay-in ceiling with a 400 mm (16 in) air space. Fig. 5-16 shows the average absorption coefficients of a sampling of 39 different acoustical tiles. The vertical lines at each frequency point indicate the spread of the coefficients for each frequency. It is interesting to note the wide variance possible with different types of tiles.

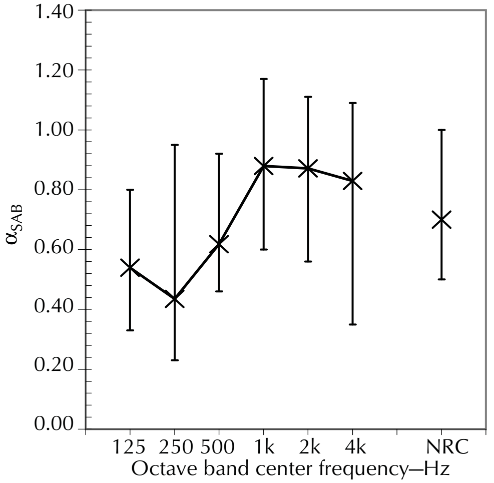


Figure 5-16. The average Sabine absorption coefficients (α_{SAB}) of 39 acoustical ceiling tiles of varying thickness, Type E400 mounting.

5.2.2.4 Spray and Trowel Applied Treatments

Some acoustical treatments can be applied by spray and/or trowel. Many are applied, finished, and detailed much like standard plaster—and are even paintable. Special bonding chemicals and processes give these types of materials their absorptive qualities. Some have a gypsum base, which can provide a look similar to normal plaster or gypsum wallboard walls. Acoustical plasters tend to provide high frequency absorption, with poor low frequency performance, especially when applied thinly (<2.5 cm thickness). Acoustical plasters can be an economical option when considering spaces that require large areas of absorption—e.g., a gymnasium ceiling. Some spray applied treatments can provide fireproofing, as well as thermal insulation. They are also popular in historical preservation applications, where the aesthetic appearance of a surface cannot be altered, but the acoustics must be improved to provide better communications in the space.

5.2.2.5 Carpet and Draperies

Carpet is a visual and comfort asset, and it is a porous absorber of sound, although principally at upper audio frequencies. Carpet is what the electrical engineer might call a low-pass filter. Because it is a high-frequency absorber, carpet should be used cautiously as a room treatment. Carpet can make a well-balanced room bass heavy because of its excessive high frequency absorption. The various types of carpet have different sound absorption characteristics. In general, sound absorption increases with pile weight and height; cut pile has greater absorption than loop pile. Pad material has a significant effect on the absorption of a carpet. Generally, the heavier the carpet pad, the more absorption. Impermeable backing should be used with care as it dramatically reduces the effect of the carpet pad and thereby reduces absorption. Due to the limited thickness of carpet, even the deepest possible pile (with the thickest possible pad) will not absorb much low-frequency sound. Fig. 5-17 shows the absorption coefficient for a typical medium-pile carpet, with and without a carpet pad.[21]

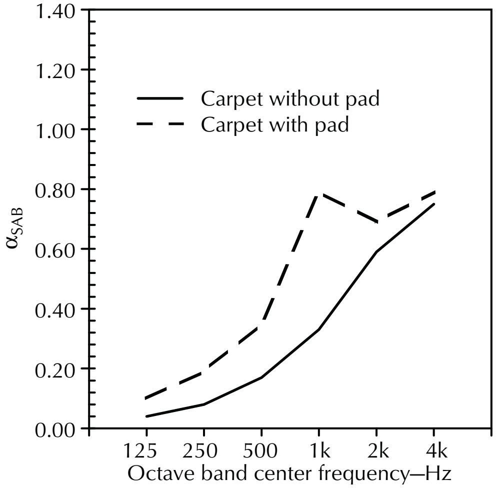


Figure 5-17. The absorption of loop pile tufted carpet (0.7 kg/m^2) with and without carpet pad (1.4 kg/m^2), Type A mounting.[21]

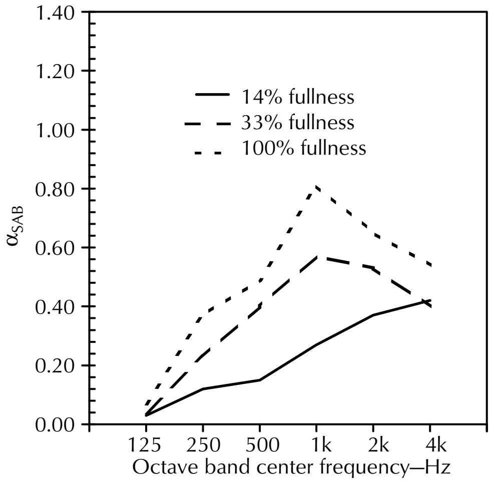


Figure 5-18. The effect of fullness on the absorption of cotton cloth curtain material, 500 g/m^2 (14.5 oz/yd^2).[22]

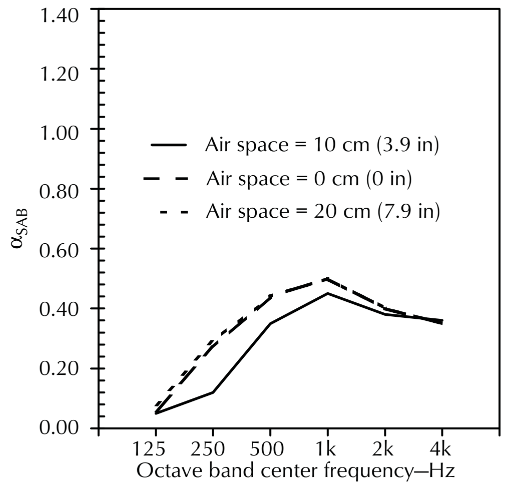


Figure 5-19. The effect of mounting over an air space on the absorption of velvet cloth curtain material, 650 g/m^2 (19 oz/yd^2).[22]

Draperies are also porous absorbers of sound. Included in the drapery category are drapes, curtains, tapestries, and other fabric wall-hangings, decorative or otherwise. Besides the type and thickness of material, the percent fullness has an effect on how well draperies absorb sound. (The percent fullness is a representation of the amount of extra material in the drapery. For example, 100% fullness would mean that a 3.0 m wide curtain actually consists of a 6.0 m wide piece of material. Similarly, 150% fullness would indicate that a 3.0 m wide drape consists of a 7.5 m wide piece of material; a 6.0 m wide piece of material being used for an 2.4 m wide drape, etc.) Fig. 5-18 shows the absorption coefficients for draperies with different percent fullness.[22] While spacing draperies from the wall does increase absorption slightly, it would not appear to be as significant as percent fullness, as indicated by Fig. 5-19.[22]

5.2.3 Discrete Absorbers

Discrete absorbers can literally be anything. Even an acoustical tile or foam panel is a discrete absorber. The absorption per unit of a tile, panel, board, person, bookshelf, equipment rack, etc., can always be determined. In the context of acoustical treatments, there are two main classes of discrete absorber that should never be ignored: people and furnishings.

5.2.3.1 People and Seats

In many large spaces, people and the seats they sit in will be the single largest acoustical treatment in the room. Any acoustical analyses of sufficiently large spaces should include people in the calculations. How the seats behave acoustically when they are empty is another important consideration. Empty wood chairs will not absorb as much as the people sitting in them. A heavily padded seat may absorb just as much sound as a seated individual. A chair that folds up when not in use may have a hard, plastic cover on the underside of the seat that will reflect sound. Perforating the cover to allow sound to pass in through the bottom of the chair when it is folded up may be a worthwhile consideration for some applications. For more information on the absorption of people, seats, and audience areas in general, refer to Section 7.3.4.4.4.

5.2.3.2 Acoustical Baffles and Banners

In very large rooms, such as domed stadiums, arenas, gymnasiums, factories, and even some houses of worship, absorbers need to be placed high on the ceiling to reduce reverberant sound. Installing spray applied acoustical treatments in such spaces is often uneconomical because it would be too labor-intensive. To solve this problem, prefabricated acoustical treatments that hang from the ceiling are often used. Acoustical baffles are typically 61 cm × 122 cm (24 in × 48 in)—or some other relatively manageable size—and are often approximately 3.8 cm (1.5 in) thick. The core material is often a rigid or semi-rigid mineral fiber, such as glass fiber, with a protective covering of polyester fabric, rip-stop nylon, or PVC. Acoustical foam panels and other porous absorption panels are often available as baffles as well. Absorption is reported as the number of sabins per baffle. Acoustical baffles are often hung vertically (perpendicular to the floor), but they can also be hung horizontal, or even at an angle. The pattern of hanging can have an effect on the overall performance of the treatments. For example, some applications will benefit more from baffles hung in two or more directions, versus simply hanging all the baffles parallel to each other in one direction. Hanging is often accomplished via factory- or user-installed grommets or hooks. Acoustical banners are simply scaled-up versions of acoustical baffles. The core absorptive material is sometimes of a slightly lower density to facilitate installing the banners so that they can be allowed to droop from a high ceiling. Sizes for banners tend to be large: 1.2 m × 15 m (4.0 ft × 50 ft) (larger sizes are not uncommon).

5.2.3.3 Other Furnishings and Objects

Anyone who has moved into a new home has experienced the absorptive power of furnishings. Rooms simply do not sound the same when they aren't filled with chairs and bookshelves and end tables and knick-knacks and so on. Even in the uncarpeted living spaces in our homes, the addition of even a small number of items can change the acoustical character of the room.

This concept was put to the test when a small room with tile floor and gypsum wallboard walls and ceilings was tested before and after the addition of two couches. The couches in question were fabric—as opposed to leather or leather substitute—and were placed roughly where they eventually wound up staying even after moving in the balance of the room's furnishings. The absorption—in sabins per couch—is shown in Fig. 5-20. (Fig. 5-20 is for illustrative purposes only—i.e., the absorption shown was not measured in a laboratory.)

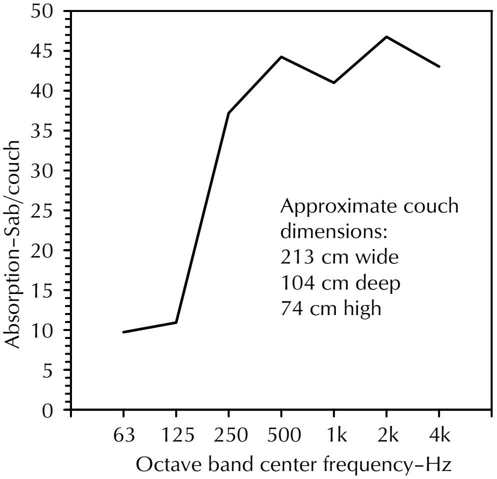


Figure 5-20. Absorption spectrum of a fabric-covered couch in a 5.6 m × 4.7 m × 2.3 m room.

5.2.4 Resonant Absorbers

In the most general sense, a resonant absorber employs the resonant properties of a material or cavity to provide absorption. Resonant absorbers are typically pressure devices, contrasted with porous absorbers that are typically velocity devices. In other words, a porous absorber placed at a point of maximum particle velocity of the sound will provide maximum absorption. A resonant absorber placed at a point of maximum particle pressure will provide maximum absorption. This can become important in applications where maximum low-frequency performance is important. A broadband porous absorber spaced away from a surface will be the most effective method of maximizing low-frequency absorption. In contrast, a resonant absorber placed on or at the surface will provide maximum low-frequency absorption.

Resonant absorbers are often described as having been tuned to address a specific frequency range. The meaning of this will become clear below from the equations involved in determining a resonant absorber's frequency of resonance. It should be noted that many versions of the equations for resonant frequency exist in the literature. Not all of these have been presented accurately and, unfortunately, some equation errors have been perpetuated. Unfortunately, calculating the resonant frequency of a resonant absorber is not straightforward. Careful research and review were undertaken for the sections below. Unless otherwise noted, the Cox and D'Antonio[2] method of utilizing the basic Helmholtz equation as the starting point for resonance calculations was implemented in the following sections.

5.2.4.1 Membrane Absorbers

Membrane absorbers—also called *panel* and *diaphragmatic* absorbers—utilize the resonant properties of a membrane to absorb sound over a narrow frequency range. Nonperforated, limp panels of wood, pressed wood fibers, plastic, or other rigid or semi-rigid material are typically employed when constructing a membrane absorber. When mounted on a solid backing, but separated from it by a constrained air space, the panel will respond to incident sound waves by vibrating. This results in a flexing of the fibers, and a certain amount of frictional loss results in absorption of some of the sound energy. The mass of the panel and the springiness of the air constitute a resonant system. In resonant systems, peak absorption occurs at the resonance frequency (f_R). The approximate f_R for a membrane absorber is given by Eq. 7-65 in Section 7.3.4.4.2. It should be emphasized that this equation yields an approximate result. Errors in calculated versus measured f_R as high as 10% have been measured.[2] Nonetheless, membrane absorbers have been successfully used to control specific resonant modes in small rooms. To control room modes, they must be placed on the appropriate surfaces at points of maximum modal pressure. (For a detailed discussion of room modes see Chapter 6.2.) Adding porous absorption, such as a mineral fiber panel (typically glass fiber or mineral wool), to the cavity dampens the resonance and effectively broadens the bandwidth or Q factor of the absorber. If the Q factor is broadened, the absorber will be somewhat effective, even if the desired frequency is not precisely attained.

Additionally, care should be taken during design and construction of membrane absorbers. Changes as small as 1 to 2 mm to, for example, the cavity depth can alter the performance significantly. Fig. 5-21 shows how the calculated resonant frequency varies with air space for various membranes. Other design tips can be found in Section 7.3.4.4.2.

Since membrane absorbers require a high level of precision to perform at the desired frequency, they are often customized for a specific application. Mass production is often uneconomical, although some companies offer membrane absorbers, one of which is the Modex Corner Bass Trap from RPG, Inc., with absorption coefficients as shown in Figure 5-22.[23]

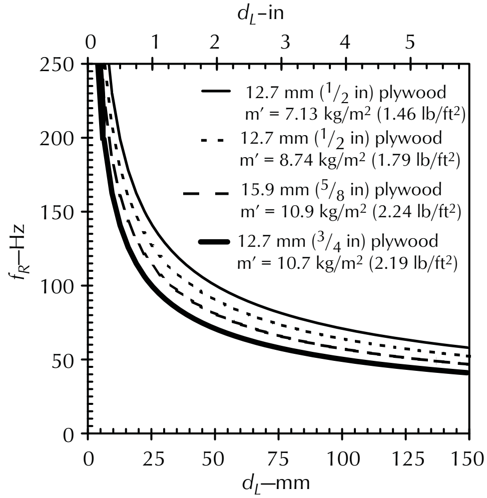


Figure 5-21. Variation of f_R versus depth of air space for membrane absorbers consisting of common building materials.

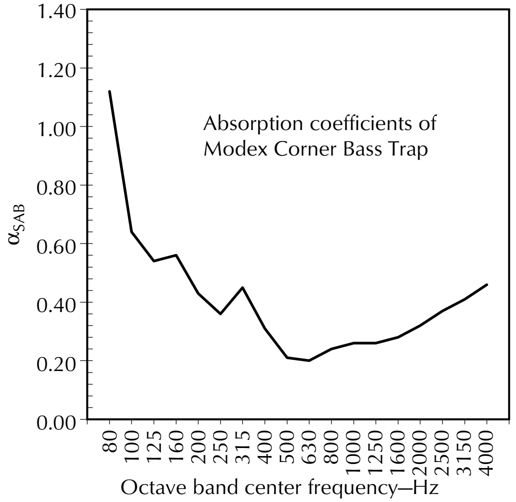
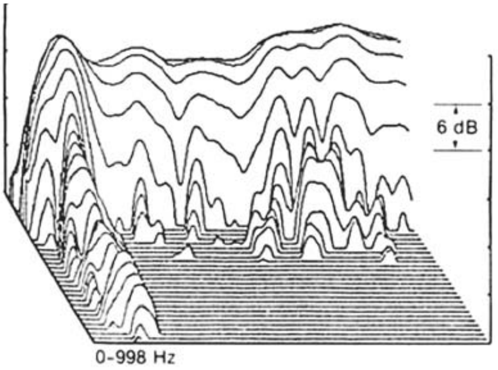


Figure 5-22. The absorption of a commercial membrane absorber, the Modex Corner Bass Trap from RPG, Inc.[23]

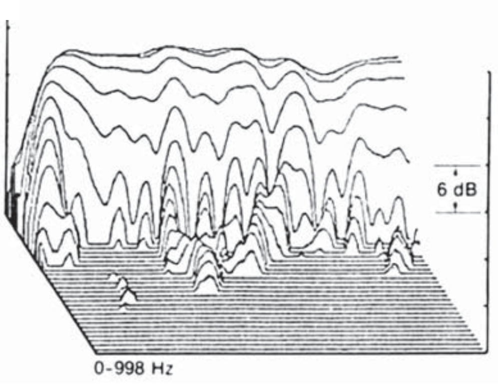
Since there have not been many mass-produced membrane absorbers, there is far less empirical test data available on membrane absorbers relative to porous absorbers. Nonetheless, some formal testing of commercially available membrane traps has been undertaken, for example, by Noy et al.[24] Results were mixed; some membrane absorbers performed as designed, others performed well (if not exactly how the designer intended), and some did not work at all.

Putting theory into practice, Fig. 5-23 shows a pair of small room response measurements before and after the addition of a membrane absorber. Frequency is plotted linearly on the x axis (horizontal) with the resonance showing up at about 140 Hz. The y axis, going into the page, is the time axis showing the decay of the room coming towards the viewer. The time span on the y axis was about 400 ms. A pair of membrane absorbers was built with f_R = 140 Hz. One was placed on the ceiling and one on a side wall.

Membrane absorbers are often inadvertently built into the structure of a room. Wall paneling, ceiling tiles, windows, coverings for orchestra pits, and even elements of furniture and millwork can all be membrane absorbers; the only question is at what frequency they resonate. Remember that everything in a room, including the room itself, has some impact on the acoustics of the room. One of the most common inadvertent membrane



A. No membrane absorber.



B. Addition of a membrane absorber.

Figure 5-23. The effect of a membrane absorber in a small room.

absorbers encountered in modern architecture is the gypsum wallboard (GWB) (drywall or sheetrock) cavity. Fortunately, the absorption of the GWB cavity can be calculated, and the calculated results have been shown to be in good agreement with laboratory measurements.[25] Section 9.2.3.1 provides discussion and calculation methods for GWB cavity absorption.

5.2.4.2 Helmholtz Resonators

The ubiquitous cola bottle may be the acoustician's most cherished conversation piece. Bottles and jugs are probably the most common everyday examples of what are referred to in acoustics as *Helmholtz resonators*. As part of his exhaustive and painstakingly detailed work in hearing, sound, and acoustics, Hermann von Helmholtz determined and documented the acoustical properties of an enclosed volume with a relatively small aperture.[26] Helmholtz resonators, as we now know them, have specialized absorptive properties for acoustical applications. At the frequency of resonance, absorption is very high. The frequency range of this absorption is very narrow—only a few Hz wide, typically. When absorptive material, such as loose mineral fiber, is used to partially fill a Helmholtz resonator, the effective frequency range is widened.

Eq. 7-69 in Section 7.3.4.4.3 can be used to calculate f_R for a Helmholtz resonator. Commercially, one of the most common products utilizing Helmholtz resonator theory is sound absorbing concrete masonry units (CMU). For example, Fig. 5-24 provides the sound absorption data for SoundBlox products available from Proudfoot.[27]

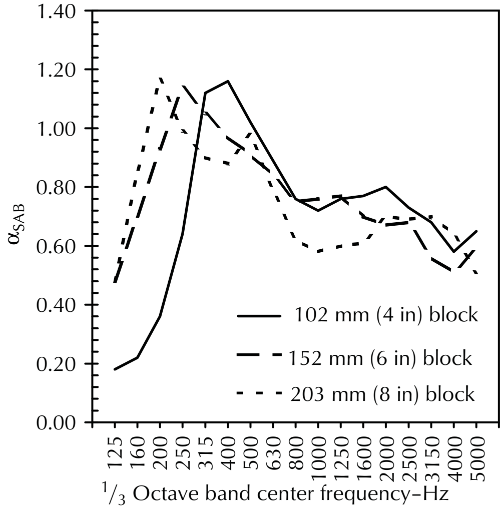


Figure 5-24. The absorption of absorbent concrete blocks from Proudfoot, Inc. Type RSC SoundBlox, utilizing Helmholtz resonance effects.[27]

5.2.4.3 Perforated Membrane Absorbers

Membrane absorbers and Helmholtz resonators are dependent on the size of the air space or cavity they contain. Turning the former into the latter can be accomplished by cutting or drilling openings in the face of the membrane. The tuned cavity of a membrane absorber then becomes the cavity of a Helmholtz resonator. When round holes are used for the openings in the face, a perforated absorber is created. To calculate the f_R for a perforated membrane absorber, first the effective thickness must be calculated. For perforated panels of having holes of diameter d and regular hole spacing S (center-to-center distance between holes), Eq. 5-2 yields the fraction of open area, ε

$$\varepsilon = \frac{\pi}{4}\left(\frac{d}{S}\right)^2 . \qquad (5\text{-}2)$$

To calculate the effective thickness for a perforated absorber, a correction factor, δ, is required. This factor is often approximated to 0.85, but can be calculated for low values of ε (typically <0.16) using

$$\delta = 0.8(1 - 1.4\sqrt{\varepsilon}) \tag{5-3}$$

Next, the effective panel thickness t' for a panel of thickness t is calculated from Eq 5-3 using δ

$$t' = t + \delta d \tag{5-4}$$

Finally, f_R for a panel over an air space of depth D is found with

$$f_R = \frac{c}{2\pi}\sqrt{\frac{\varepsilon}{t'D}} \tag{5-5}$$

Care should be taken to be consistent with units. For example, if inches are used to calculate ε, etc., c (the speed of sound) should be in inches per second.

The f_R of perforated absorbers is generally adjusted by changing ε. Increasing ε (larger holes, smaller spacing, or both), decreasing D, or using thinner panels will all increase the f_R. The f_R can be lowered by decreasing ε, by increasing D, or by using thicker panels. The f_R from Eq. 5-5 is not exact, but is close enough for use in the design stage. The air space is often partially or completely filled with porous absorption. The only drawback to this is that absorptive material in contact with the perforated panel can reduce the absorber's effectiveness.

One of the more obvious perforated membranes that can be used is common pegboard. Standard pegboard tends to create an absorber with an f_R in the 250–500 Hz range, as shown in Fig. 5-25.[18] Since perforated absorbers are often considered for low frequency control, it is not uncommon to fabricate customized perforated boards. For example, a hardboard membrane with $d = 6.4$ mm (¼ in), $S = 102$ mm (4 in), $t = 3.2$ mm (⅛ in), and $D = 51$ mm (2 in), a perforated absorber tuned to roughly 125 to 150 Hz can be created. The absorption coefficients of such an absorber with 96 kg/m³ (6.0 lb/ft³) of glass fiber filling the air space are shown in Fig. 5-26.

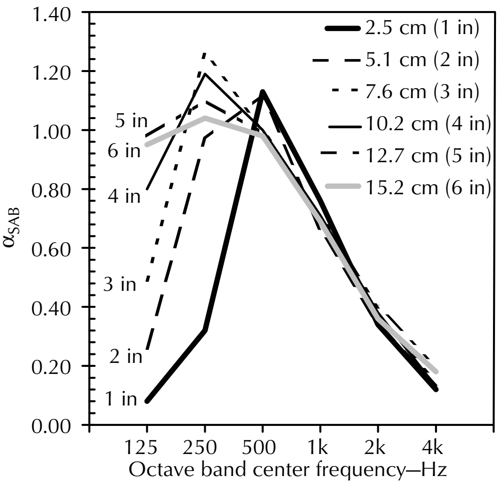


Figure 5-25. The effect of pegboard facing on the absorption of different thicknesses of Owens Corning 703 glass fiber boards, Type A mounting.[18]

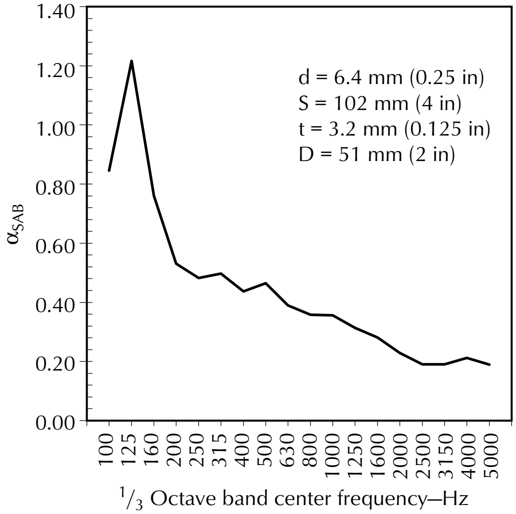


Figure 5-26. The absorption of a perforated membrane absorber "tuned" to 125–150 Hz, Type A mounting, cavity filled with 96 kg/m³ (6 lb/ft³) glass fiber.

Microperforated materials are one of the most recent developments in the area of acoustical treatments. Extremely thin materials with tiny perforations (<<1 mm) are stretched over an air space and absorption occurs by means of boundary layer effects.[2] Because they are so thin, microperforated absorbers can be fabricated from visually transparent material. RPG offers the ClearSorber, which can be installed by stretching it over glass without significantly altering the light throughput. The absorption coefficients, dependent on the depth of air space between the microperforated foil and the glass, are shown in Fig. 5-27.[28]

5.2.4.4 Slat Absorbers

Helmholtz resonators can also be constructed by using spaced slats over an air space (with or without absorptive

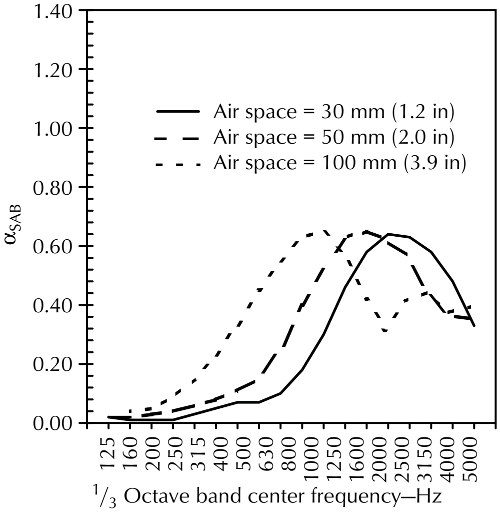


Figure 5-27. The effect of depth of air space on the absorption of a 105 μm thick microperforated absorber, the RPG ClearSorber Foil from RPG, Inc.[28]

fill). The air mass in the slots between the slats reacts with the springiness of the air in the cavity to form a resonant system, much like that of the perforated panel type. In fact, Eq. 5-5 is again used to calculate the f_R for a slat absorber, but with the following equations for ε, δ, and t':

$$\varepsilon = \frac{r}{w+r} \tag{5-6}$$

$$\delta = -\frac{1}{\pi}\ln\left[\sin\left(\frac{1}{2}\pi\varepsilon\right)\right] \tag{5-7}$$

$$t' = t + 2\delta r \tag{5-8}$$

where,
r is the slot width,
w is the slat width.

While δ is often approximated to a value near 0.6, it is not difficult to calculate. As with perforated absorbers, the above will yield approximate results for the f_R of a slat absorber, which will be fine for most design applications.

In a practical sense, the absorption curve can be broadened by using a variable depth for the air space. Another approach is using slots of different widths. In the structure of Fig. 5-28, both variable air space depth and variable slot width could be used. Porous absorptive fill is shown at the back of the cavity, removed from the slats. This gives a sharper absorption than if the absorbent is in contact with the slats. It should be noted that, all other factors remaining the same, randomly placed slats (yielding randomly sized slots) will lower the overall absorption, while bandwidth is increased.[29]

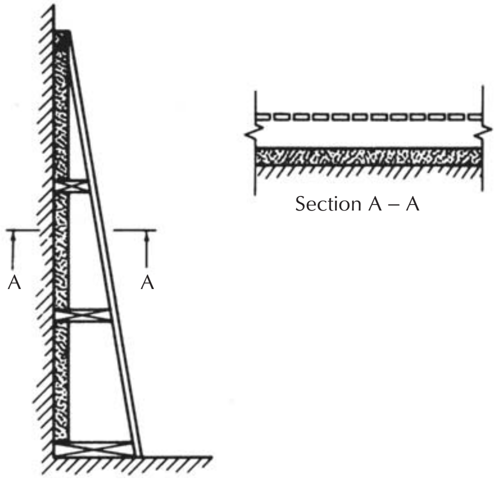


Figure 5-28. A slat absorber with varying depth to widen the effective frequency range of absorption.

5.2.5 Bass Traps

Bass trap lore pervades the professional audio industry, particular in the recording industry. Literature on the subject, however, is scarce. The term has become a catchphrase often used by acoustical product manufacturers to include a large variety of acoustical products, often including what are simply broadband absorbers. Very few bass traps are actually effective at absorbing bass. It is simply quite difficult to absorb sounds with wavelengths at or approaching 56 ft (17 m). To most effectively absorb a given frequency at any angle of incidence, including normal incidence the absorbing material should be at least $^1/_{10}$ and, ideally, ¼ of the wavelength of the lowest frequency of interest. For 20 Hz, this means a depth between 1.7 m (minimum) and 4.3 m (ideal)! It is relatively rare to find someone who is willing to build a device that large to trap bass. This may be necessary in the design of very large rooms, like concert halls, but it is probably more interesting to ask what crime has the 20 Hz committed that it needs to be trapped? As we shall see in the next chapter, the low end performance of small rooms can be reliably predicted from a study of the distribution of room modes. If the modes are distributed properly, trapping may not be needed. On the other hand, imagine that the modes are not distributed properly and the goal is to fix a problem room. If there is enough space to build a bass trap that is large enough to have an impact on wavelengths that large, most likely one could move a wall, improving, if not optimizing, the modal distribution and obviating the need for trapping.

Nonetheless, the preceding sections have provided many examples of products that could be designed to trap bass without taking up much space. Additionally, there are good broadband absorbers on the market that extend down into the low-frequency region. The practical limits of size and mounting usually lead to a natural cutoff between 50 and 100 Hz for many broadband absorbers. These products exhibit performance that is highly dependent on placement, especially in small rooms.

5.2.6 Applications of Absorption

In large rooms (see Chapter 6 for the definition of large vs small rooms) where there is a statistically reverberant sound field, absorption can be used to actually modify the reverberant field, and the results are predictable and fairly straightforward. The whole concept of reverberation time (RT—discussed in detail in Chapter 6) is a statistical one that is based on the assumption of uniform distribution of energy in the room and complete randomness in the direction of sound propagation. In large rooms, both conditions can prevail.

In small rooms—particularly at low frequencies—the direction of propagation is by no means random. Because of this, the propriety of applying the common RT equations to small rooms is questioned. For small rooms (nonreverberant spaces), absorption is useful for control of discrete reflections from surfaces located in the near field of the source and listener. In rooms the size of the average recording studio or home theater, a true reverberant field cannot be found. In such small rooms, the common RT equations cannot be used reliably. Further, predicting or trying to measure RT in small spaces where a reverberant field cannot be supported is typically less useful relative to other analysis techniques. The results of RT measurements in small rooms will neither show RT in the true sense, nor will they reveal much of value regarding the behavior of sound in a small room relative to the time domain. Typically, it is more useful to examine the behavior of sound in the time domain in more detail in small rooms. Determining the presence of reflections (wanted or unwanted), the amplitude of those reflections, and their direction of arrival at listening positions is typically a better approach. For low frequencies, Chapter 5 provides some small room analysis techniques that are more beneficial than the measurement of RT.

5.2.6.1 Use of Absorption in Reverberant Spaces

In large rooms, the common RT equations can be used with reasonable confidence. When absorptive treatment is not uniformly distributed throughout the space, the Sabine formula is typically avoided in lieu of other formulas. RT is covered in detail in Chapter 6.

In reverberant spaces, the selection of absorbers can be based on the absorption data collected in accordance with ASTM C423, as described in Section 5.2.1.1. Care should be taken, however, to somehow account for effects not directly evident from laboratory measurement methods. As one example, consider a fabric-wrapped, mineral fiber panel that is tested in a Type A mounting configuration. The test specimen is placed in the reverberation chamber directly against a hard (typically solid concrete) surface, often the floor. The absorption coefficients then represent only the absorption provided by the panels. In practice, panels such as these might be directly applied to a GWB surface having absorption characteristics of its own that are significantly different than the solid concrete floor of a reverberation chamber! Applying an absorptive panel to a GWB wall or ceiling not only changes the acoustical behavior of the GWB surface (by changing the mass), but the panel itself will not absorb as measured in the lab, because of the mounting, the size of the room relative to the laboratory test chamber, the proportion of absorptive material relative to the total surface area of the room, etc. This is one example of why the predictive modeling for the acoustics of large spaces can be—like many aspects of acoustics—as much art as science. All acousticians are likely to have methods they use to account for idiosyncrasies that can neither be measured in a laboratory, nor modeled by a computer.

In addition, it is generally agreed among acousticians that reverberation time is no longer considered the single most important parameter in music hall and large auditorium acoustics. Reverberation time is understood to be one of several important criteria of acoustical quality of such halls. Equal or greater stress is now placed on, for example, the ratio of early arriving energy to total sound energy, the presence of lateral reflections, the timing of the arrival of various groups of reflections, and other parameters discussed in detail in Chapter 6.

5.2.6.2 Use of Absorption in Nonreverberant Spaces

In rooms where there is not enough volume or a long enough mean free path to allow a statistical reverberant field to develop, one must view the use of absorption in a somewhat different manner. As alluded to previously, the common RT equations will not work satisfactorily in these spaces.

Absorption is often used in small rooms to control discrete reflections or to change the way the room feels. Contrary to popular belief, the impression of liveness or deadness is not based on the length of the reverberation time. Rather, it is based on the ratio of direct to reflected sound and on the timing of the early reflected sound field, especially in the first 20 ms or so. Adjusting the acoustics of nonreverberant spaces (sometimes referred to as *tuning the room*) involves manipulating discrete reflections.

To determine the suitability of a particular absorber, the acoustician needs a direct measurement of the reflected energy off of the product. In small rooms—where a significant portion of the spectrum is below f_1 (see Chapter 6)—field measurement of absorption, such as the techniques and methods presented in Section 5.2.1.3, might be more appropriate in determining the applicability of a particular absorber to a small room application.

5.2.7 Subjective Impact of Absorption

Sometimes it is useful to consider the extremes. It is interesting to note that rooms with no absorption and rooms with total absorption represent the most acoustically hostile spaces imaginable. At one extreme, there is the absorption-free space, also known as the reverberation chamber. A good real-world example of this is a racquetball court. As anyone who has played racquetball can readily attest to, racquetball courts are not acoustically friendly! At the other extreme is the anechoic chamber. This is a room that is totally absorptive and totally quiet. Since an anechoic chamber has no reflected sound and is isolated from sounds from the outside, a good real-world example of this is the desert. Standing in a part of the desert free of reflective surfaces, such as buildings and mountains, located many kilometers from any noise sources, such as highways and people, at a time when there is no wind, the complete lack of sound would be comparable to what one would experience in an anechoic chamber. It is difficult to describe just how disorienting spending time in either of these chambers can be. Neither the reverberation chamber nor the anechoic chamber is a place where a musician would want to spend much time, let alone perform!

The use of absorption has a powerful impact on the subjective performance of a room. If too much absorption is used, the room will feel too dead—i.e., too much like an anechoic chamber. If too little absorption is used, the room will feel uncomfortably live—i.e., too much like a reverberation chamber. Additionally, the absorption of any material or device is frequency-dependent; absorbers act like filters to the reflected sound. Some energy is turned into heat, but other frequencies are reflected back into the room. Choosing an absorber that has a particularly nonlinear response can result in rooms that just plain sound strange.

More often than not, the best approach is a combination of absorbers. For example, large spaces that already have the seats, people, and carpet as absorbers may benefit from a combination of membrane absorbers and porous absorbers. In a small room, some porous absorbers mixed with some Helmholtz resonators might provide the best sound for the room. Both of these are examples of the artistic (the aural aesthetic) being equally applied with the science (the acoustic physics).

Experience is important when considering the application of absorption and—more importantly—when considering what a particular application will sound like. The savvy acoustician will realize the aural differences between a small room treated with 5.1 cm (2 in) acoustical foam and a room treated with 2.5 cm (1 in) glass fiber panels. On paper, these materials are quite similar (compare Figs. 5-6 and 5-14). However, the knowledge that a room treated with 9.3 m^2 (100 ft^2) of foam generally sounds darker than a room treated with the same area of 96 kg/m^3 fabric-wrapped, glass fiber boards comes only with experience. Likewise, the knowledge that a room treated with a slotted concrete block wall will sound much different than the same room with a GWB wall that is treated with several well-placed perforated absorbers (even though RT predictions for each scenario come out to be approximately the same) comes only with experience.

5.2.8 Absorption and Absorption Coefficients of Common Building Materials and Absorbers

Table 5-3 gives the absorption coefficients of various popular building materials.

5.3 Acoustical Diffusion

Compared to acoustical absorption, the science of acoustical diffusion is relatively new. The oft-cited starting point for much of the science of modern diffusion is the work of Manfred Schroeder. In fact, acoustically diffusive treatments that are designed using one of the various numerical methods that will be discussed below are often referred to generically as Schroeder diffusers. In the most

Table 5-3. Absorption Data for Common Building Materials and Acoustical Treatments (All Materials Type A Mounting Unless Noted Otherwise)

Material		125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	Source
Walls & Ceilings								
Brick, unpainted		0.03	0.03	0.03	0.04	0.05	0.07	Ref. 21
Brick, painted		0.01	0.01	0.02	0.02	0.02	0.03	Ref. 21
Concrete block, unpainted		0.36	0.44	0.31	0.29	0.39	0.25	Ref. 21
Concrete block, painted		0.10	0.05	0.06	0.07	0.09	0.08	Ref. 21
One layer 13 mm (½″) GWB, Mounted on each side of 90 mm (3.5″) metal studs. No cavity insulation		0.26	0.10	0.05	0.07	0.04	0.05	Ref. 25
Two layers 13 mm (½″) GWB. Mounted on each side of 90 mm (3.5″) metal studs. No cavity insulation		0.15	0.08	0.06	0.07	0.07	0.05	Ref. 25
One layer 13 mm (½″) GWB. Mounted on each side of 90 mm (3.5″) metal studs. With glass fiber cavity insulation		0.14	0.06	0.09	0.09	0.06	0.05	Ref. 25
One layer 13 mm (½″) GWB. Mounted on one side of 90 mm (3.5″) metal studs. With or without cavity insulation		0.12	0.10	0.05	0.05	0.04	0.05	Ref. 25
Plaster over tile or brick, smooth finish		0.01	0.02	0.02	0.03	0.04	0.05	Ref. 21
Plaster on lath, rough finish		0.14	0.10	0.06	0.05	0.04	0.03	Ref. 21
Plaster on lath, smooth finish		0.14	0.10	0.06	0.04	0.04	0.03	Ref. 21
Floors								
Heavy carpet without pad		0.02	0.06	0.14	0.37	0.60	0.65	Ref. 21
Heavy carpet with pad		0.08	0.24	0.57	0.69	0.71	0.73	Ref. 21
Concrete or terrazzo		0.01	0.01	0.02	0.02	0.02	0.02	Ref. 21
Linoleum, rubber, cork tile on concrete		0.02	0.03	0.03	0.03	0.03	0.02	Ref. 21
Parquet over concrete		0.04	0.04	0.07	0.06	0.06	0.07	Ref. 21
Marble or glazed tile		0.01	0.01	0.01	0.01	0.02	0.02	Ref. 21
Other								
Ordinary window glass		0.35	0.25	0.18	0.12	0.07	0.04	Ref. 21
Double glazing (1.4–1.6 cm thick)		0.10	0.07	0.05	0.03	0.02	0.02	Ref. 2
Water surface		0.01	0.01	0.01	0.02	0.02	0.03	Ref. 21
Acoustical Treatments	**Fig.**							
2.5 cm (1″) Owens Corning 701	4-4	0.17	0.33	0.64	0.83	0.90	0.92	Ref. 15
2.5 cm (1″) Owens Corning 703	4-4, 4-6	0.11	0.28	0.68	0.90	0.93	0.96	Ref. 15
2.5 cm (1″) Owens Corning 705	4-4	0.02	0.27	0.63	0.85	0.93	0.95	Ref. 15
10.2 cm (4″) Owens Corning 701	4-5	0.73	1.29	1.22	1.06	1.00	0.97	Ref. 15
10.2 cm (4″) Owens Corning 703	4-5, 4-6	0.84	1.24	1.24	1.08	1.00	0.97	Ref. 15
10.2 cm (4″) Owens Corning 705	4-5	0.75	1.19	1.17	1.05	0.97	0.98	Ref. 15
5.1 cm (2″) Owens Corning 703	4-6	0.17	0.86	1.14	1.07	1.02	0.98	Ref. 15
7.6 cm (3″) Owens Corning 703	4-6	0.53	1.19	1.21	1.08	1.01	1.04	Ref. 15
12.7 cm (5″) Owens Corning 703	4-6	0.95	1.16	1.12	1.03	1.04	1.06	Ref. 15
15.2 cm (6″) Owens Corning 703	4-6	1.09	1.15	1.13	1.05	1.04	1.04	Ref. 15
2.5 cm (1″) Owens Corning. Linear Glass Cloth Board. No airspace	4-7	0.05	0.22	0.60	0.92	0.98	0.95	Ref. 15
2.5 cm (1″) Owens Corning. Linear Glass Cloth Board. 2.5 cm (1″) airspace	4-7	0.04	0.26	0.78	1.01	1.02	0.98	Ref. 15
2.5 cm (1″) Owens Corning. Linear Glass Cloth Board. 5.1 cm (2″) airspace	4-7	0.17	0.40	0.94	1.05	0.97	0.99	Ref. 15
2.5 cm (1″) Owens Corning. Linear Glass Cloth Board. 7.6 cm (3″) airspace	4-7	0.19	0.83	1.03	1.04	0.92	1.00	Ref. 15
2.5 cm (1″) Owens Corning. Linear Glass Cloth Board. 12.7 cm (5″) airspace	4-7	0.41	0.73	1.02	0.98	0.94	0.97	Ref. 15

Table 5-3. Absorption Data for Common Building Materials and Acoustical Treatments (All Materials Type A Mounting Unless Noted Otherwise) (Continued)

2.5 cm (1″) Roxul RockBoard 40	4-9	0.07	0.32	0.77	1.04	1.05	1.05	Ref. 17
3.8 cm (1½″) Roxul RockBoard 40	4-9	0.18	0.48	0.96	1.09	1.05	1.05	Ref. 17
5.1 cm (2″) Roxul RockBoard 40	4-9, 4-10	0.26	0.68	1.12	1.10	1.03	1.04	Ref. 17
Material	**Fig.**	**125 Hz**	**250 Hz**	**500 Hz**	**1000 Hz**	**2000 Hz**	**4000 Hz**	**Source**
7.6 cm (3″) Roxul RockBoard 40	4-9	0.63	0.95	1.14	1.01	1.03	1.04	Ref. 17
10.2 cm (4″) Roxul RockBoard 40	4-9	1.03	1.07	1.12	1.04	1.07	1.08	Ref. 17
5.1 cm (2″) Roxul RockBoard 35	4-10	0.26	0.68	1.14	1.13	1.06	1.07	Ref. 17
5.1 cm (2″) Roxul RockBoard 60	4-10	0.32	0.81	1.06	1.02	0.99	1.04	Ref. 17
5.1 cm (2″) Roxul RockBoard 80	4-10	0.43	0.78	0.90	0.97	0.97	1.00	Ref. 17
2.5 cm (1″) Tectum Wall Panel	4-11	0.06	0.13	0.24	0.45	0.82	0.64	Ref. 18
3.8 cm (1½″) Tectum Wall Panel	4-11	0.07	0.22	0.48	0.82	0.64	0.96	Ref. 18
5.1 cm (2″) Tectum Wall Panel	4-11	0.15	0.26	0.62	0.94	0.62	0.92	Ref. 18
5.1 cm (2″) Auralex Studiofoam Wedge	4-13, 4-14	0.11	0.30	0.91	1.05	0.99	1.00	Ref. 19
5.1 cm (2″) Auralex Studiofoam Pyramid	4-13	0.13	0.18	0.57	0.96	1.03	0.98	Ref. 19
5.1 cm (2″) Auralex Studiofoam Metro	4-13	0.13	0.23	0.68	0.93	0.91	0.89	Ref. 19
5.1 cm (2″) Auralex Sonomatt	4-13	0.13	0.27	0.62	0.92	1.02	1.02	Ref. 19
5.1 cm (2″) Auralex Sonoflat	4-13	0.16	0.46	0.99	1.12	1.14	1.13	Ref. 19
5.1 cm (1″) Auralex Studiofoam Wedge	4-14	0.10	0.13	0.30	0.68	0.94	1.00	Ref. 19
5.1 cm (3″) Auralex Studiofoam Wedge	4-14	0.23	0.49	1.06	1.04	0.96	1.05	Ref. 19
5.1 cm (4″) Auralex Studiofoam Wedge	4-14	0.31	0.85	1.25	1.14	1.06	1.09	Ref. 19
2.5 cm (1″) SONEXmini	4-15	0.11	0.17	0.40	0.72	0.79	0.91	Ref. 20
3.8 cm (1½″) SONEXmini	4-15	0.14	0.21	0.61	0.80	0.89	0.92	Ref. 20
5.1 cm (2″) SONEXclassic	4-15	0.05	0.31	0.81	1.01	0.99	0.95	Ref. 20
7.6 cm (3″) SONEXone	4-15	0.09	0.68	1.20	1.18	1.12	1.05	Ref. 20
Pegboard with 6.4 mm (¼″) holes on 2.5 cm (1″) centers over 2.5 cm (1″) thick Owens-Corning 703	4-25	0.08	0.32	1.13	0.76	0.34	0.12	Ref. 15
Pegboard with 6.4 mm (¼″) holes on 2.5 cm (1″) centers over 5.1 cm (2″) thick Owens-Corning 703	4-25	0.26	0.97	1.12	0.66	0.34	0.14	Ref. 15
Pegboard with 6.4 mm (¼″) holes on 2.5 cm (1″) centers over 7.6 cm (3″) thick Owens-Corning 703	4-25	0.49	1.26	1.00	0.69	0.37	0.15	Ref. 15
Pegboard with 6.4 mm (¼″) holes on 2.5 cm (1″ centers over 10.2 cm (4″) thick Owens-Corning 703	4-25	0.80	1.19	1.00	0.71	0.38	0.13	Ref. 15
Pegboard with 6.4 mm (¼″) holes on 2.5 cm (1″) centers over 12.7 cm (5″) thick Owens-Corning 703	4-25	0.98	1.10	0.99	0.71	0.40	0.20	Ref. 15
Pegboard with 6.4 mm (¼″) holes on 2.5 cm (1″) centers over 15.2 cm (6″) thick Owens-Corning 703	4-25	0.95	1.04	0.98	0.69	0.36	0.18	Ref. 15

basic sense, diffusion can be thought of as a special form of reflection. Materials that have surface irregularities on the order of the wavelengths of the impinging sound waves will exhibit diffusive properties. Ideally, a diffuser will redirect the incident acoustical energy equally in all directions and over a wide range of frequencies. However, it is often impractical to construct a device that can diffuse effectively over the entire audible frequency range. Most acoustical diffuser products are designed to work well over a specific range of frequencies, typically between 2 and 4 octaves above roughly 500 Hz. Of course, just as with absorbers, one must be concerned with the performance of a diffuser.

5.3.1 Diffuser Testing: Diffusion, Scattering, and Coefficients

The performance of a diffuser can be expressed as the amount of diffusion and as the amount of scattering provided by a surface. While there is still some disagreement as to diffusion nomenclature, Cox and D'Antonio have attempted to establish a distinction between diffusion and scattering, particularly as it relates to the coefficients that are used to quantify diffuser performance.[2] In general, diffusion and the diffusion coefficients relate to the uniformity of the diffuse sound field created by a diffuser. This is most easily explained by looking at polar plots of the diffuse sound field from a diffuser. The less lobing there is in a diffusion polar plot, the more diffusion and the higher the diffusion coefficients.

Scattering and the scattering coefficients relate to the amount of energy that is not reflected in a specular manner. The term specular here denotes the direction of reflection one would expect if the sound were reflecting off a hard flat surface. For example, most high frequency sounds reflect from a hard flat surface at the same angle as the incident sound, Fig. 5-1. This is referred to as specular reflection by Cox and D'Antonio. The more sound that is reflected in a nonspecular manner, the higher the scattering. Therefore, a simple angled wall can provide high scattering but low diffusion, since the reflected sound will still form a lobe, but not in a specular direction. It should be noted that a significant amount of absorption makes it difficult to measure scattering. This makes sense since an absorber does not allow for a high level of specular reflection; absorption can be mistaken for scattering.

The standardized methods for measuring diffusion and scattering coefficients are AES-4id-2001 and ISO 15664, respectively.[30,31] The AES method provides guidelines for measuring the performance of diffusive surfaces and reporting the diffusion coefficients. These guidelines will allow diffusers of different designs to be objectively compared. The results cannot, however, be incorporated into acoustical modeling programs. For that, the scattering coefficients must be used when measured in accordance with the ISO method.

Both the AES and ISO methods are relatively new; the AES method was formalized in 2001 and the ISO method was published in 2003. Because of this, none of the independent acoustical test laboratories in North America are equipped to perform the AES diffusion test and, as of this writing, only one laboratory in North America is equipped to perform the ISO scattering test. Because of this, diffusion and scattering coefficients for surfaces and treatments (diffusers or otherwise) are difficult to find. Indeed, because so little testing is being performed on diffusers, there is some degree of confusion in the industry as to what diffusion and scattering coefficients actually mean in subjective terms. For example, what does a diffusion (or scattering) coefficient of 0.84 at 2500 Hz sound like? There is no denying that the information is useful and that objective quantification of diffusers is necessary. However, comparing diffusion or scattering coefficients for different materials would be a theoretical exercise at best. The process is further complicated by the fact that commercial diffusers vary dramatically in shape and style; each manufacturer claims some degree of superiority because of some unique application of some innovative mathematics.

Nonetheless, there are diffusion and scattering coefficients available in the literature (Cox and D'Antonio offer a significant amount of laboratory measured diffusion and scattering coefficients[2]), and some manufacturers have begun pursuing independent tests of their diffusive offerings. Fig. 5-29 pictures examples of various commercial diffusers. The next few decades will be a very exciting time for diffusion, particularly if more independent acoustical laboratories begin to offer AES and/or ISO testing services.

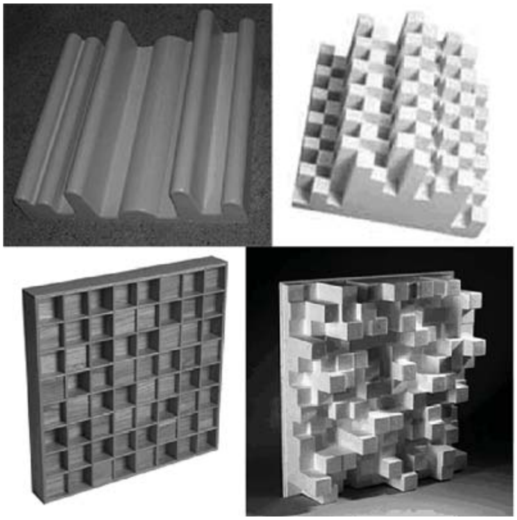


Figure 5-29. Various commercial diffusers.

5.3.2 Mathematical (Numerical) Diffusers

The quadratic residue diffuser (QRD) is one form of a family of diffusers known as reflection phase gratings, or more generally, mathematical or numerical diffusers. Numerical diffusers, such as the QRD, are based on the pioneering work of Manfred Schroeder.[32] Numerical diffusers consist of a periodic series of slots or wells of equal width, with the depth determined by a number theory sequence. The depth sequence is developed via Eq. 5-9

$$well\ depth\ factor = n^2 \bmod p \qquad (5\text{-}9)$$

where,
p is a prime number,
n is an integer ≥ 0.

The "mod" in Eq. 5-9 refers to modulo, which is a number theory mathematical process whereby the first number, in this case the square of n, is divided by the second number, in this case p, and the remainder is equal to the well depth factor. For example, if $n = 5$ and $p = 7$, Eq. 5-9 becomes

$$Well\ depth\ factor = 25 mod\ 7$$

25 divided by 7 = 3 with a remainder of 4.
Therefore:

$$\begin{aligned} Well\ depth\ factor &= 25 mod\ 7 \\ &= 4 \end{aligned}$$

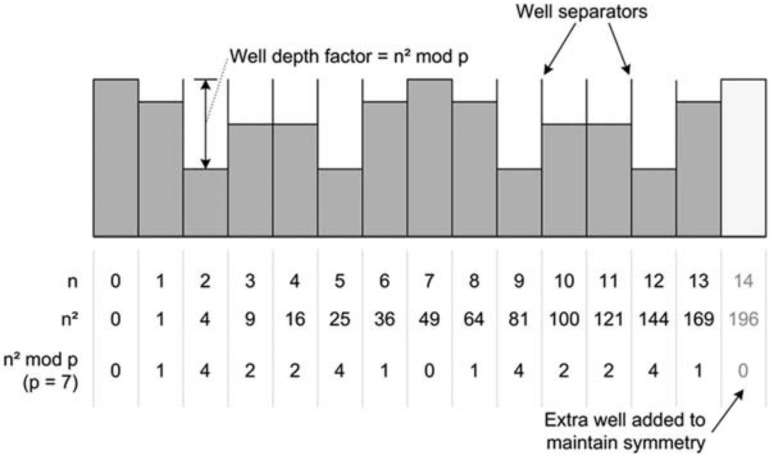
In a similar manner, the well depth factors for all the other wells are obtained, as shown in Fig. 5-30. Two complete periods—plus an extra well added to maintain symmetry—are shown in Fig. 5-30 for a $p = 7$ QRD. Usually the wells are separated by thin, rigid separators (but not always). An important feature of QRD is symmetry. This allows them to be manufactured and utilized in multiple modular forms.

D'Antonio and Konnert have outlined the theory and application of reflection phase-grating diffusers.[32] The maximum frequency for effective diffusion is determined by the width of the wells; the minimum frequency for effective diffusion is determined by the depth of the well. Commercial diffusers built on these principles are available from a variety of manufacturers. RPG, Inc. and its founder, Dr. Peter D'Antonio, have done pioneering work in the area of diffusion, particularly with respect to Schroeder diffusion and, more recently, with state-of-the-art diffusive surfaces that are customized for an application through the use of special computer models and algorithms.

Numerical diffusers such as QRDs can be 1D or 2D in terms of the diffused sound pattern. QRDs consisting of a series of vertical or horizontal wells will diffuse sound in the horizontal or vertical directions, respectively. In other words, if the wells run in the ceiling-floor direction, diffusion will occur laterally, from side to side, and the resulting diffusion pattern would resemble a cylinder. (Incident sound parallel to the wells will be reflected more than diffused.) More complex numerical diffusers employ sequences of wells—often elevated blocks or square-shaped depressions—that vary in depth (or height) both horizontally and vertically. Incident sound striking these devices will be diffused in a spherical pattern.

5.3.3 Random Diffusion

Besides numerical diffusers, diffusion can also result from the randomization of a surface. In theory, these surfaces cannot provide ideal diffusion. However, listening to the results after treating the surfaces of a room with random diffusers would indicate that, subjectively, they perform quite well. Since any randomization of a surface breaks up specular reflection to some degree, this is not unexpected. The only limitation will be the frequency range of significant diffusion. The rules for well width and depth discussed previously would still apply, albeit in a general sense since the diffusers will not have been designed using a formal number theory algorithm. The benefit of random diffusion is that everyday materials and objects can provide significant diffusion. For example, bookshelves, media storage shelves, decorative trim or plasterwork, furnishings, fixtures, and other decorations can all provide some diffusion. This can be


Well separators
Well depth factor = n² mod p
Extra well added to maintain symmetry

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
n^2	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196
n^2 mod p ($p = 7$)	0	1	4	2	2	4	1	0	1	4	2	2	4	1	0

Figure 5-30. A profile of two periods (with one extra well to maintain symmetry) of a QRD of prime number $p = 7$.

particularly helpful when budget is a concern since diffusive treatments tend to cost twice to ten times more than absorptive treatments on a per square foot basis.

5.3.4 Applications of Diffusion

Like absorption, major differences in the size of the space being treated need to be considered when applying diffusion principles. Diffusion tends to provide the most per-dollar-spent benefit in large spaces. Diffusion tends to happen in the far field. In the near field of a diffuser, the scattering effects are typically less pronounced and can actually be less subjectively pleasing than a flat reflective surface in some applications.

For large spaces, diffusion is often employed on the ceiling or rear wall of a space. This provides distribution of the sound energy in the room that envelops a listener. There will be a noticeable (and measurable) reduction in RT, but the reduction is not nearly as severe as it would be with a similar area of absorptive treatment in the same space. Additionally, a typical listener tends not to notice that RT has been reduced, but rather that the decay has been smoothed out, intelligibility has improved, and the room generally sounds better after diffusion has been appropriately applied.

For small spaces, the decision to use diffusion is more challenging. D'Antonio and Cox suggest that the full benefit of diffusers is realized when the listener is a minimum working distance of about 3 m (10 ft) away from the diffusive surface.[34] This tends to be a good rule-of-thumb from which to start. This rule provides something of a size threshold that must be exceeded to get the most value from the application of diffusers. When small rooms are being treated, it is not uncommon to be able to get far more value from absorption relative to the same area of diffusion.

The most common applications of diffusion in small rooms, such as recording studios and residential theaters or listening rooms, is (like larger spaces) on the rear wall and ceiling. The application of diffusion to the rear wall was particularly popular in the heyday of Live End/Dead End (LEDE) recording studio design. While LEDE is still a popular approach, it has been replaced with the more general reflection-free zone (RFZ) approach to control room design. Regardless of the design method, diffusion can be useful for removing reflective artifacts from a small room without making the room too dead. It can also be used to great effect, for example, on the ceiling of a home theater. A diffusive, rather than absorptive, ceiling can sometimes provide a better feel to a home theater, regardless of the ceiling height.

5.4 Reflection and Other Forms of Sound Redirection

In addition to diffusion, sound can be redirected by controlled reflection, diffraction, or refraction. Diffraction takes place when sound bends around an object, such as when a passing train is audible behind a wall. The low frequencies from the train rumble have large wavelengths relative to the height of the wall, allowing them to bend over the top of it. This can come into play indoors when sound bends around office partitions, podiums, or other common obstacles.

Refraction is the only form of acoustic redirection that does not involve some sort of object. Acoustic refraction is the bending of a sound wave caused by changes in the velocity of sound through a medium. Refraction is often thought of as an optical phenomenon; however, acoustic refraction occurs when there are temperature gradients in a room. Because the speed of sound is dependent on the temperature of the air, when an acoustic wave passes through a temperature gradient, it will bend toward the cooler air. This can occur indoors in large rooms when cooler air from air-conditioning vents located on the ceiling blows into a room with warmer air below; the sound will bend upwards until the temperature reaches equilibrium. Even in recording studios, a heating vent blowing warm air over one loudspeaker in a stereo pair can skew the sound towards the other loudspeaker and wreak havoc on the stereo image!

Finally, everything in the room, including the room itself, reflects sound in some way, even the absorbers. One could look at an absorber as an inefficient reflector. When an item is small with respect to the wavelength of sound impinging upon it, it will have little effect on that sound. A foot stool placed in front of a woofer will have little effect on the 1.7 m wavelength of 200 Hz, the 4.4 m wavelength of 100 Hz, or the 6.9 m wavelength of 50 Hz. The wave will diffract around it and continue on its merry way. The wavelength of 8 kHz is 4.3 cm. Just about anything that is placed in front of a tweeter that is reproducing 8 kHz and up can effectively block or redirect the sound.

Although it may seem odd, reflection is a very useful form of acoustic treatment, especially in concert halls. Adding reflections to the direct sound is what makes concert halls what they are. However, adding reflections to a monitoring environment can impede critical analysis by coloring the sound coming from the speakers. It is not a question of reflections being good or bad; rather, the designer must decide to include or exclude reflections based on the use of the facility and the desired outcome.

For more information on the design of usefully reflective surfaces for large spaces, see Chapter 7, especially Section 7.3.4.

5.5 Electronic Treatments

Absorptive acoustical treatments effectively add damping to the room. The reduction of decay is generally the goal. It is a myth that electronics can be used in place of absorptive acoustical treatments. There is no electronic device that can be inserted into the signal path that will prevent sound from a loudspeaker from reflecting off the surfaces of the room. Nonetheless, since the beginning of the electroacoustic era, devices such as electronic absorbers and room equalizers have been proposed. Not all of these are without merit. As early as 1953, Olson and May proposed an electronic sound absorber consisting of a microphone, amplifier, and loudspeaker.[35] Over a short distance from the microphone, the device could be tuned to achieve as much as 10 to 20 dB of practical attenuation over a 1 to 2 octave range of low frequencies. Olson and May proposed that their electronic sound absorbers could be used to reduce noise at the ears of airline passengers and factory workers. Unfortunately, the ineffectiveness of this type of absorber over larger distances made it impractical for use in architectural applications. The concept, however, paved the way for future developments.

The invention of the parametric equalizer (PEQ) brought a new wave of hope for electroacoustical treatments. Unfortunately, the insertion of a PEQ into the signal chain, even to reduce narrowband problems in small rooms, usually caused more harm than good. Because of the variability of the sound pressure distribution in a small room, the desired effect of the PEQ was usually limited to a small area of the room. Additionally, phase anomalies usually made the treatment sound unnatural. The use of a PEQ to tune a recording studio control room, for example, came and went quickly and for good reason.

The age of digital signal processing, combined with the availability of high-quality audio equipment to a wider range of users, such as home theater owners, ushered in a new hope for electroacoustical treatments. The most recent devices, while sometimes referred to as room equalization (as in previous decades), are often referred to as digital room correction, or DRC. The most important improvement of these devices over their analog ancestors is their ability to address sound problems occurring in the time/phase domains. The latest in DRC systems are able to address minimum-phase problems, such as axial room modes (see Chapter 6). These problems often manifest themselves not as amplitude problems (which are what would be addressed in the use of analog equalizers), but as decay problems. More modern DRC systems, such as those developed by Wilson et al, that incorporate the latest in digital signal processing, can now actually add the damping that is required to address minimum-phase low-frequency problems.[36] Additionally, many DRC systems require that the room response be measured at multiple listening locations in the room so that algorithms can be used to determine corrections that can benefit a larger area of the room.

The same advances in signal processing have also brought about wider applications for the original electronic sound absorber of Olson and May. Bag End has developed the E-Trap, an electronic bass trap that offers the ability to add significant and measurable damping at two different low frequencies.[37]

While DRC devices and electronic traps offer much in the way of being able to actually address the problems with the loudspeaker-room interface, they cannot be expected to be more than electronic tweaks. They cannot replace a good acoustical room design with proper incorporation of nonelectronic treatments. They can provide some damping, particularly in the lowest octave or two where in many rooms it is often impractical—if not impossible—to incorporate porous or resonant absorbers.

5.6 Acoustical Treatments and Life Safety

The most important consideration when selecting acoustical treatments is safety. Most often, common sense should prevail. For example, asbestos acoustical treatments—which were quite popular several decades ago—should be avoided because of the inherent health risks associated with handling asbestos materials and breathing its fibers. Acoustical treatments will have to meet any applicable building codes and safety standards to be used in a particular facility. Specific installations may also dictate that specific materials be avoided because of allergies or special use of the facility—e.g., health care or correctional facilities. Since many acoustical treatments will be hung from walls and ceilings, only the manufacturer-approved mounting methods should be used to prevent injury from falling objects. The two most common health and safety concerns for acoustical treatment materials are flammability and breathability.

Acoustical treatments must not only meet the applicable fire safety codes, but, in general, should not be flammable. The flammability of an interior finish such as an acoustical treatment is typically tested in accordance with the ASTM E84 standard to measure flammability.[38] The results of the ASTM E84 test are a flame spread index and a smoke developed index. Building codes further classify materials according to the test results. International Building Code (IBC) classifications are as follows:[39]

- Class A: Flame spread index = 0–25, smoke developed index = 0–450.
- Class B: Flame spread index = 26–75, smoke developed index = 0–450.
- Class C: Flame spread index = 76–200, smoke developed index = 0–450.

Materials to be used as interior finishes, such as acoustical treatments, are often tested in accordance with ASTM E84, with test results provided in manufacturer literature. The ASTM E84 test results and corresponding IBC classifications for some typical acoustical materials are summarized in Table 5-4.

In general, most acoustical materials are Class A materials. Some acoustical foam treatments, as well as some acoustical treatments made of wood, are Class B. Care should be taken that any acoustical treatments made of foam or wood have been tested and that the manufacturer can provide proof of testing. It should also be noted that some jurisdictions require that acoustical foam materials be subjected to more stringent flammability requirements, such as the NFPA 286 test method.[40]

With regards to breathability, precautions should be taken if the acoustical treatment material contains fibers that could be respiratory or skin irritants. The fibers of many common acoustical treatments, such as glass fiber and mineral wool panels, are respiratory and skin irritants, but are harmless once the treatments have been installed in their final configuration, usually with a fabric or other material encasing the fibrous board. Nonetheless, precautions such as wearing gloves and breathing masks should be taken when handling the raw materials or when installing the panels. Additionally, damaged panels should be repaired or replaced in order to minimize the possibility of fibers becoming airborne.

Some facilities may have additional safety requirements. Some health care facilities may disallow porous materials of any kind to minimize the possibility of, for example, mold or bacterial growth. Clean room facilities may also prohibit the use of porous materials on the grounds of minimizing the introduction of airborne particles. Correctional facilities will often prohibit any materials that can be burned (including some fire-resistant materials) and securing acoustical treatment panels to walls or ceilings without any removable mechanical fasteners, such as screwed, rivets, bolts, etc. Still other facilities may have safety requirements based on, for example, the heat produced by a piece of machinery, the chemicals involved in a manufacturing process, and so on. The applicable laws, codes, and regulations—including rules imposed by the end user— should always be consulted prior to the purchase, construction, and installation of acoustical treatments.

5.7 Acoustical Treatments and the Environment

Acoustical treatments should be selected with an appropriate level of environmental awareness. Depending on the application, selection could include not only what the material itself is made of, but also how it is made, how it is transported to the facility, and how it will be disposed of should it be replaced sometime in the future. Many acoustical treatments, such as those consisting of natural wood or cotton fibers, can contribute to Leadership in Energy and Environmental Design (LEED) certification. Unlike audio electronics where overseas manufacturing has become the norm, acoustical treatments are often manufactured and fabricated locally, thus saving on the financial and environmental costs of transportation.

Even acoustical treatments such as polyurethane foam panels, which are a byproduct of the petroleum refining process and can involve the use of carbon dioxide (a greenhouse gas) in the manufacturing process, are becoming more environmentally friendly. For example, one manufacturer of acoustical foam products, Auralex Acoustics, Inc., has begun using soy components in their polyurethane products, thereby reducing the use of carbon-rich petroleum components by as much as 60%.

The best possible environmentally friendly approach to the use of acoustical treatments is to limit their use. The better a facility can be designed from the beginning, the fewer specialty acoustical treatment materials will be required. Rooms from recording studios to cathedrals that are designed with acoustics in mind from the beginning generally require relatively fewer specialty acoustical treatments. Acoustical treatments are difficult to avoid completely; almost every space where the production or reproduction of sound takes place, or where the ability to communicate is tantamount, will require some acoustical treatment. Nonetheless, the most conservative approach to facility design should ensure that only those acoustical treatments that are absolutely necessary are implemented in the final construction.

Table 5-4. Typical ASTM E84 Test Results and Corresponding IBC Classifications for Common Acoustical Treatments

Acoustical Treatment	Flame Spread Index	Smoke Developed Index	IBC Class	Comments
Glass Fiber panels	15	0	A	Unfaced material
Mineral wool panels	5	10	A	Unfaced material
Wood fiber panels	0	0	A	Unfaced, treated material
Cotton fiber panels	10	20	A	Unfaced, treated material
Acoustical foam panels Polyurethane	35	350	B	Unfaced, treated material NFPA 286 test may also be required
Acoustical foam panels Melamine	5	50	A	Unfaced material NFPA 286 test may also be required
Acoustical plaster	0	0	A	Unfaced material
Acoustical Diffusers Polystyrene	15	145	A	Treated material NFPA 286 test may also be required
Acoustical Diffusers Wood	25	450	A	Treated material

References

1. Don Davis, *If Bad Sound Were Fatal, Audio Would Be the Leading Cause of Death*, Bloomington, IN: 1st Books, 2004.
2. Trevor J. Cox and Peter D’Antonio, *Acoustic Absorbers and Diffusers: Theory, Design and Application,* London and New York: Spon Press, 2004.
3. W.C. Sabine, *Collected Papers in Acoustics*, Cambridge: Harvard University Press, 1922.
4. Leo L. Beranek, *Acoustics*, Woodbury, NY: Acoustical Society of America, 1993.
5. *Standard Test Method for Sound Absorption and Sound Absorption Coefficients by the Reverberation Room Method.* ASTM C423-07a.
6. *Acoustics — Measurement of Sound Absorption in a Reverberation Room*. ISO 354:2003.
7. *Standard Practices for Mounting Test Specimens during Sound Absorption Tests*. ASTM E795-05.
8. Peter D’Antonio, “Bass Management,” *Diffuse Reflections* newsletter, Vol. 10, issue 1, 2005.
9. Peter D’Antonio, “Bass Management,” *Diffuse Reflections* newsletter, Vol. 10, issue 2, 2005.
10. Albert London, “The Determination of Reverberant Sound Absorption Coefficients from Acoustic Impedance Measurements,” *Journal of the Acoustical Society of America*, Vol. 22, no. 2, pp. 263–269, March 1950.
11. Brad Nelson, “Single Reflection Analysis in Determining Absorption Coefficients,” *Synergetic Audio Concepts Newsletter,* Vol. 8, no. 1, p. 14, Fall 1980. Other comments in Vol. 26, no. 1, Winter 1998.
12. *Acoustics — Sound Absorbers for Use in Buildings — Rating of Sound Absorption*. ISO 11654:1997.
13. Pranab Saha, “Do We Know What We Are Doing?” *Sound and Vibration*, Vol. 35, no. 4, April 2001.
14. John S. Parkinson, “Area and Pattern Effects in the Measurement of Sound Absorption,” *Journal of the Acoustical Society of America*, Vol. 2, no. 1, pp. 112–122, July 1930.
15. *Noise Control Manual*, sixth edition, Publication no. 5-BMG-8277-G, Toledo, OH: Owens Corning Fiberglass Corp. April 1986.
16. F.P. Mechel, “Design Charts for Sound Absorber Layers,” *Journal of the Acoustical Society of America*, Vol. 83, no. 3, pp. 1002–1013, March 1988.
17. Various data sheets available from Roxul®, Inc., http://www.roxul.com/sw34127.asp, accessed March 9, 2008.
18. Data available from Tectum, Inc., http://www.tectum.com/technicaldata/Wall%20Panel%20Test%20Data.pdf, accessed March 9, 2008.
19. Data available from Auralex Acoustics, Inc., http://www.auralex.com/testdata/test/mastertable.pdf, accessed March 9, 2008.
20. Various data sheets available from Pinta Acoustic, Inc., http://www.pinta-acoustic.com/literature.php, accessed March 9, 2008.
21. Cyril M. Harris (Editor in Chief), *Handbook of Acoustical Measurements and Noise Control,* Woodbury, NY: Acoustical Society of America, 1998.

22. V. S. Mankovsky, *Acoustics of Studios and Auditoria*, New York: Focal Press, 1971.
23. Data available from RPG, Inc., http://www.rpginc.com/products/clearsorberf/index.htm, accessed March 28, 2008.
24. Dirk Noy, Gabriel Hauser, and John Storyk, "Low Frequency Absorbers—Applications and Comparisons," *Presented at the 115th Convention of the Audio Engineering Society*, Preprint No. 5944, October 2003.
25. J.S. Bradley, "Sound Absorption of Gypsum Board Cavity Walls," *Journal of the Audio Engineering Society*, Vol. 45, no. 4, pp. 253–259, April 1997.
26. Hermann L.F. Helmholtz, *On the Sensations of Tone*, Mineola, NY: Dover, 1954. (An unaltered, unabridged re-publication of the Ellis translation of the second edition, originally published by Longmans & Co. in 1885.)
27. Data available from Proudfoot, Inc., http://www.soundblox.com/soundblox/absorptioncoefficient.php, accessed March 9, 2008.
28. Data available from RPG, Inc., http://www.rpginc.com/products/clearsorberf/index.htm, accessed March 28, 2008.
29. M. Rettinger, "Low-Frequency Slot Absorbers," *db The Sound Engineering Magazine,* Vol. 10, no. 6, pp. 40–43, June 1976.
30. *AES Information Document for Room Acoustics and Sound Reinforcement System—Characterization and Measurement of Surface Scattering Uniformity*, Audio Engineering Society, 2001.
31. *Acoustics — Sound-Scattering Properties of Surfaces — Part 1: Measurement of the Random-Incidence Scattering Coefficient in a Reverberation Room*. ISO 17497-1:2004.
32. M.R. Schroeder, "Diffuse Sound Reflection by Maximum-Length Sequences," *Journal of the Acoustical Society of America*, Vol. 57, no. 1, pp. 149–150, January 1975.
33. Peter D. Antonio and John H. Konnert, "The Reflection Phase Grating Diffusor: Design Theory and Application," *Journal of the Audio Engineering Society*, Vol. 32, no. 4, pp 228–238, April 1984.
34. Peter D'Antonio and Trevor Cox, "Technical Bulletin on the Application of Diffusion in Critical Listening Rooms," Upper Marlboro, MD: RPG, Inc., 2003.
35. Harry F. Olson and Everett G. May, "Electronic Sound Absorber," *Journal of the Acoustical Society of America*, Vol. 25, no. 6, pp. 1130–1136, November 1953.
36. Rhonda J. Wilson, Michael D. Capp, and J. Robert Stuart, "The Loudspeaker-Room Interface—Controlling Excitation of Room Modes," *Presented at the 23rd International Conference of the Audio Engineering Society*, May 2003.
37. Reza Kashani and James Wischmeyer, "Electronic Bass Trap," *Presented at the 117th Convention of the Audio Engineering Society*, Preprint No. 6277, October 2004.
38. *Standard Test Method for Surface Burning Characteristics of Building Materials*. ASTM E84-01.
39. *2006 International Building Code®*, International Code Council, 2006.
40. *Standard Methods of Fire Tests for Evaluating Contribution of Wall and Ceiling Interior Finish to Room Fire Growth*. NFPA 286, 2006 edition.

Bibliography

Trevor J. Cox and Peter D'Antonio, *Acoustic Absorbers and Diffusers: Theory, Design and Application*, London and New York: Spon Press, 2004.

W.C. Sabine, *Collected Papers in Acoustics*, Cambridge: Harvard University Press, 1922.

Heinrich Kuttruff, *Room Acoustics*, New York: Spon Press, 2000.

C. Zwikker and C.W. Kosten, *Sound Absorbing Materials*, New York: Elsevier Publishing Company, Inc., 1949.

Uno Ingard, *Noise Reduction Technology, Part 1*, *ABSORBERS*, Kittery Point, ME: http://www.lulu.com/, 1999.

Chapter 6

Small Room Acoustics

by Doug Jones

6.1 Introduction 127
6.2 Room Modes 127
6.3 Modal Room Resonances 128
  6.3.1 Comparison of Modal Potency 131
  6.3.2 Modal Bandwidth 131
6.4 Criteria for Evaluating Room Modes 132
6.5 Modes in Nonrectangular Rooms 133
6.6 Summation of Modal Effect 134
6.7 The Question of Reverberation 135
  6.7.1 Room Shape 138
  6.7.2 Reflection Manipulation 138
  6.7.3 Comb Filters 139
6.8 Small Room Design Factors 140
  6.8.1 The Control Room Controversy 140
  6.8.2 Precision Listening Rooms 140
  6.8.3 Rooms for Microphones 141
6.9 Rooms for Entertainment 142
References 143

This page intentionally left blank

6.1 Introduction

The acoustics of small rooms is dominated by modes, shape, and reflection management. Acousticians who build large rooms are frequently frustrated with small room design because few of the intellectual tools of the trade that work in large rooms can be applied to small rooms. Getting small rooms to sound right involves art and science. The science part is mostly straightforward. The creative part is quite subjective and a great sounding small room can be just as elusive as a great sounding concert hall.

6.2 Room Modes

A room mode is a phenomenon that occurs whenever sound travels between two reflecting surfaces where the distance between the surfaces is such that the impinging wave reflects back on itself creating a standing wave. The distribution of modes determines the low frequency performance of a small room. Consider a sound source *S* emitting a sinusoidal signal between two isolated reflecting surfaces as in Fig. 6-1. Starting at a very low frequency, the frequency of the oscillator driving the source is slowly increased. When a frequency of $f_0 = 1130/2L$ (in feet) is reached, a so-called standing-wave condition is set up. Consider what is happening at the boundary. Particle velocity must be zero at the wall surface but wherever particle velocity is zero, pressure is at maximum level. The wave is reflected back out of polarity with itself, that is to say that the reflection is delayed by ½ of the period. This results in a cancellation that will occur exactly midpoint between the reflecting surfaces. If the walls are not perfect reflectors, losses at the walls will affect the heights of the maxima and the depths of the minima. In Fig. 6-1 reflected waves traveling to the left and reflected waves traveling to the right interfere, constructively in some places, destructively in others. This effect can be readily verified with a sound level meter which will show maximum sound pressures near the walls and a distinct null midway between the walls.

As the frequency of the source is increased, the initial standing-wave condition ceases, but at a frequency of $2f_0$ another standing wave appears with two nulls and a pressure maximum midway between the walls. Other standing waves can be set up by exciting the space between the walls at whole number multiples of f_0. These are called axial modes as they occur along the axis of the two parallel walls.

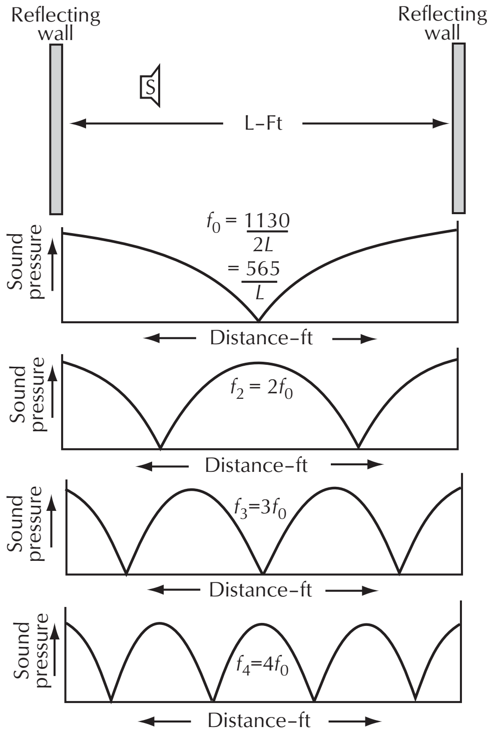


Figure 6-1. The simplest form of room resonance can be illustrated by two isolated, parallel, reflecting wall surfaces.

The two walls of Fig. 6-1 can be considered the east and west walls of a room. The effect of adding two more pairs of parallel walls to enclose the room is that of adding two more axial standing-wave systems, one along the east-west axis and the other along the vertical axis. In addition to the two axial systems that are set up, there will be a standing wave associated with two times the path length that involves all four surfaces. These modes are called tangential modes, Fig. 6-2. Most rooms will have six boundaries and there are modes that involve all six surfaces as well, Fig. 6-3. These modes are called oblique modes.

In 1896 Lord Rayleigh showed that the air enclosed in a rectangular room has an infinite number of normal or natural modes of vibration. The frequencies at which these modes occur are given by the following equation:[1]

$$f = \frac{c}{2}\sqrt{\left(\frac{p}{L}\right)^2 + \left(\frac{q}{W}\right)^2 + \left(\frac{r}{H}\right)^2} \quad (6\text{-}1)$$

where,

c is the speed of sound, 1130 ft/s (or 344 m/s),
L is the length of the room in feet (or meters),
W is the width of the room in feet (or meters),
H is the height of the room in feet (or meters),
p, q, and *r* are the integers 0, 1, 2, 3, 4, and so on.

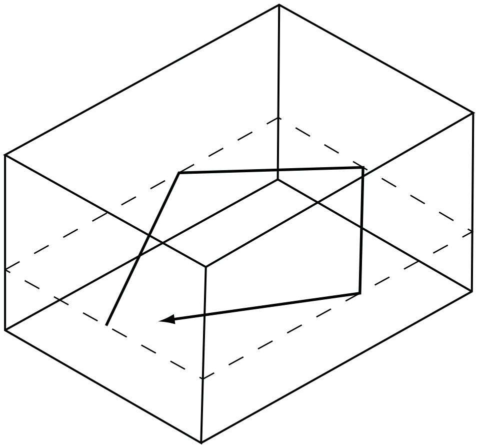


Figure 6-2. Tangential room modes.

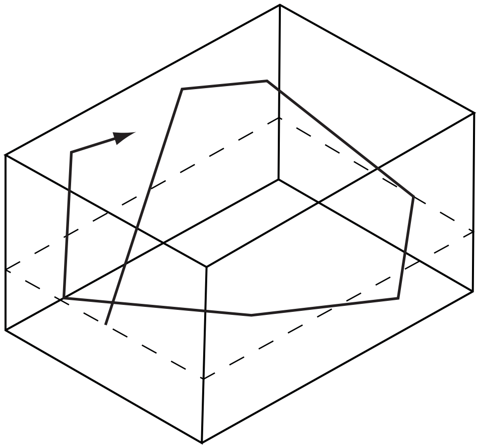


Figure 6-3. Oblique room modes.

If we consider only the length of the room, we set q and r to zero, their terms drop out, and we are left with

$$\begin{aligned} f &= \frac{c}{2} - \sqrt{\left(\frac{1}{L}\right)^2} \\ &= \left(\frac{c}{2} - \frac{1}{L}\right) \\ &= \frac{1130}{2L} \qquad (6\text{-}2) \\ &= \frac{565}{L} \text{ in feet} \\ &= \frac{172}{L} \text{ in meters} \end{aligned}$$

which looks familiar because it is the f_0 frequency of Fig. 6-1. Thus, if we set $p = 1$ the equation gives us f_0. When $p = 2$ we get $2f_0$, with $p = 3$, $3f_0$ and so on. Eq. 6-1 covers the simple axial mode case, but it also presents us with the opportunity of studying forms of resonances other than the axial modes.

Eq. 6-1 is a 3D statement based on the orientation of our room on the x, y, and z axes, as shown in Fig. 6-4. The floor of the room is taken as the x plane, and the height is along the z axis. To apply Eq. 6-5 in an orderly fashion, it is necessary to adhere to standard terminology. As stated, p, q, and r may take on values of zero or any whole number. The values of p, q, and r in the standard order are thus used to describe any mode. Remember that:

- p is associated with length L.
- q is associated with width W.
- r is associated with height H.

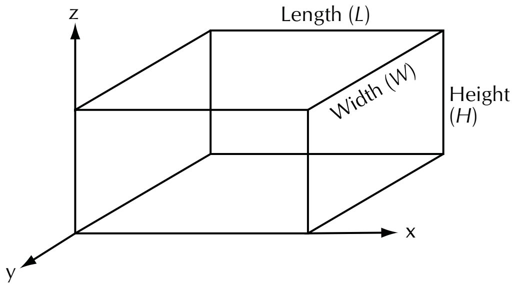


Figure 6-4. The floor of the rectangular room under study is taken to be in the *xy* plane and the height along the *z* axis.

We can describe the four modes of Fig. 6-1 as 1,0,0; 2,0,0; 3,0,0; and 4,0,0. Any mode can be described by three digits. For example, 0,1,0 is the first-order width mode, and 0,0,2 is the second-order vertical mode of the room. Two zeros in a mode designation mean that it is an *axial mode*. One zero means that the mode involves two pairs of surfaces and is called a *tangential mode*. If there are no zeros in the mode designation, all three pairs of room surfaces are involved, and it is called an *oblique mode*.

6.3 Modal Room Resonances

In order to better understand how to evaluate the distribution of room modes, we calculate the modal frequencies for three rooms. Let us first consider a room with dimensions that *are not* recommended for a sound room. Consider a room with the dimensions of 12 ft long, 12 ft wide by 12 ft high (3.66 m × 3.66 m × 3.66 m), a perfect cube. For the purposes of this exercise, let us assume that all the reflecting surfaces are solid and massive. Using Eq. 6-1 to calculate only the axial modes (for now) we see a fundamental mode at 565/12 or 47.08 Hz. If we continue the series, looking at the 1,0,0; 2,0,0; 3,0,0…10,0,0 modes we see the results in Table 6-1.

Table 6-1. Modal Frequencies for a 12 Ft Cube Room (Axial Only)

Length	Modes	Length	Modes
47.08	1,0,0	282.50	6,0,0
94.17	2,0,0	329.58	7,0,0
141.25	3,0,0	376.67	8,0,0
188.33	4,0,0	423.75	9,0,0
235.42	5,0,0	470.83	10,0,0

Before we continue the calculation, let us examine what this table is indicating. The frequencies listed are those and only those that are supported by these two walls; that is to say there will be some resonance at these frequencies but at no others. When the source is cut off, the energy stored in a mode decays logarithmically. The actual rate of decay is determined by the type of mode and the absorptive characteristics of whatever surfaces are involved with that mode. An observer in this situation, making a sound with frequency content that includes 141 Hz, may hear a slight increase in amplitude depending on the location in the room. The observer will also hear a slightly longer decay at 141 Hz. At 155 Hz, for example, there will be no support or resonance anywhere between these two surfaces. The decay will be virtually instantaneous as there is no resonant system to store the energy. Of course, in a cube the modes supported by the other dimensions (0,1,0; 0,2,0; 0,3,0 … 0,10,0 and 0,0,1; 0,0,2; 0,0,3… 0,0,10) will all be identical, Table 6-2.

Table 6-2. Axial Modes in a Cube Supported in Each Dimension

Length	Modes	Width	Modes	Height	Modes
47.08	1,0,0	47.08	0,1,0	47.08	0,0,1
94.17	2,0,0	94.17	0,2,0	94.17	0,0,2
141.25	3,0,0	141.25	0,3,0	141.25	0,0,3
188.33	4,0,0	188.33	0,4,0	188.33	0,0,4
235.42	5,0,0	235.42	0,5,0	235.42	0,0,5
282.50	6,0,0	282.50	0,6,0	282.50	0,0,6
329.58	7,0,0	329.58	0,7,0	329.58	0,0,7
376.67	8,0,0	376.67	0,8,0	376.67	0,0,8
423.75	9,0,0	423.75	0,9,0	423.75	0,0,9
470.83	10,0,0	470.83	0,10,0	470.83	0,0,10

In the cube, all three sets of surfaces are supporting the same frequencies and no others. Talking in such a room is like singing in the shower. The shower stall supports some frequencies, but not others. You tend to sing at those frequencies because the longer decay at those frequencies adds a sense of fullness to the sound. Table 6-2 can be made more useful by listing all the modes in order to better examine the relationship between them. Table 6-3 is such a listing. In this table, we have included the spacing in Hz between a mode and the one previous to it.

Table 6-3. All of the Axial Modes of the Cube in Table 6-1 and Table 6-2

Frequency	Modes	Spacing	Frequency	Modes	Spacing
47.08	1,0,0		282.50	6,0,0	47.08
47.08	0,1,0	0.00	282.50	0,6,0	0.00
47.08	0,0,1	0.00	282.50	0,0,6	0.00
94.17	2,0,0	47.08	329.58	7,0,0	47.08
94.17	0,2,0	0.00	329.58	0,7,0	0.00
94.17	0,0,2	0.00	329.58	0,0,7	0.00
141.25	3,0,0	47.08	376.67	8,0,0	47.08
141.25	0,3,0	0.00	376.67	0,8,0	0.00
141.25	0,0,3	0.00	376.67	0,0,8	0.00
188.33	4,0,0	47.08	423.75	9,0,0	47.08
188.33	0,4,0	0.00	423.75	0,9,0	0.00
188.33	0,0,4	0.00	423.75	0,0,9	0.00
235.42	5,0,0	47.08	470.83	10,0,0	47.08
235.42	0,5,0	0.00	470.83	0,10,0	0.00
235.42	0,0,5	0.00	470.83	0,0,10	0.00

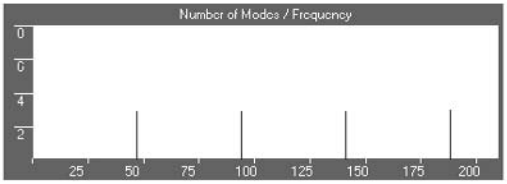


Figure 6-5. Number of axial modes and frequencies for a cube room. From AcousticX.

We can clearly see the triple modes that occur at every axial modal frequency, and there is 47 Hz (equal to f_0) between each cluster. The space between each cluster is important because if a cluster of modes or even a single mode is separated by more than about 20 Hz from it nearest neighbor, it will be quite audible as there is no masking from nearby modes. Consider another room that does not have a good set of dimensions for a sound room, but represents a typical room size because of standard building materials. Our next test room is 16 ft long by 12 ft wide with a ceiling height of 8 ft (4.88 m × 3.66 m × 2.44 m). Table 6-4 shows the length, width, and height modes, and Table 6-5 shows the same data sorted into order according to frequency.

Table 6-4. Room Modes of a Rectangular Room 16 Ft × 12 Ft × 8 Ft

Length	Modes	Width	Mode	Height	Modes
35.31	1,0,0	47.08	0,1,0	70.63	0,0,1
70.63	2,0,0	94.17	0,2,0	141.25	0,0,2
105.94	3,0,0	141.25	0,3,0	211.88	0,0,3
141.25	4,0,0	188.33	0,4,0	282.50	0,0,4
176.56	5,0,0	235.42	0,5,0	353.13	0,0,5
211.88	6,0,0	282.50	0,6,0	423.75	0,0,6
247.19	7,0,0	329.58	0,7,0	494.38	0,0,7
282.50	8,0,0	376.67	0,8,0	565.00	0,0,8
317.81	9,0,0	423.75	0,9,0	635.63	0,0,9
353.13	10,0,0	470.83	0,10,0	706.25	0,0,10

Table 6-5. Modes of a 16 Ft × 12 Ft × 8 Ft Room Sorted by Frequency

Frequency	Modes	Spacing	Frequency	Modes	Spacing
35.31	1,0,0		282.50	8,0,0	35.31
47.08	0,1,0	11.77	282.50	0,6,0	0.00
70.63	2,0,0	23.54	282.50	0,0,4	0.00
70.63	0,0,1	0.00	317.81	9,0,0	35.31
94.17	0,2,0	23.54	329.58	0,7,0	11.77
105.94	3,0,0	11.77	353.13	10,0,0	23.54
141.25	4,0,0	35.31	353.13	0,0,5	0.00
141.25	0,3,0	0.00	376.67	0,8,0	23.54
141.25	0,0,2	0.00	423.75	0,9,0	47.08
176.56	5,0,0	35.31	423.75	0,0,6	0.00
188.33	0,4,0	11.77	470.83	0,10,0	47.08
211.88	6,0,0	23.54	494.38	0,0,7	23.54
211.88	0,0,3	0.00	565.00	0,0,8	70.63
235.42	0,5,0	23.54	635.63	0,0,9	70.63
247.19	7,0,0	11.77	706.25	0,0,10	70.63

- If we examine the data in Fig. 6-6 we see that there are some frequencies which are supported by only one dimension. 35 Hz, for example, is only supported by the 16 ft (4.88 m) dimension. Other frequencies, like 70 Hz, occur twice and are supported by length and height. Still others like 141 Hz occur three times and are supported by all three dimensions. In the nomograph of Fig. 6-6, the height of the line indicates the magnitude of the mode. This room is clearly better than a cube, but it is far from ideal as there are many frequencies which will stand out. 70 Hz, 141 Hz, 211 Hz, 282 Hz, and 253 Hz are all going to be problem frequencies in this room.

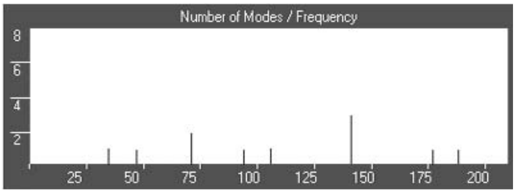


Figure 6-6. Number of modes and frequencies for a room 16 ft × 12 ft × 8 ft. From AcousticX.

Now consider a room that has dimensions that might be well suited for an audio room; 23 ft long by 17 ft wide by 9 ft high ceiling (7 m × 5.18 m × 2.74 m). The sorted data and nomograph are shown in Table 6-6 and Fig. 6-7.

The data in Fig. 6-7 look quite different from the data in Fig. 6-5 and Fig. 6-6. There are no instances where all three dimensions support the same frequency. There is also a reasonably good distribution of modes across the spectrum. There are a few places where the difference between the modes is quite small, like the space between the 4,0,0 and the 0,3,0. These three rooms, the cube, the room with dimensions determined by the builder, and the last room demonstrate the first important principle when dealing with room modes.

Table 6-6. Data for a Room 23 Ft × 17 Ft × 9 Ft

Frequency	Modes	Spacing	Frequency	Modes	Spacing
24.57	1,0,0		196.52	8,0,0	8.19
33.24	0,1,0	8.67	199.41	0,6,0	2.89
49.13	2,0,0	15.90	221.09	9,0,0	21.68
62.78	0,0,1	13.65	232.65	0,7,0	11.56
66.47	0,2,0	3.69	245.65	10,0,0	13.01
73.70	3,0,0	7.23	251.11	0,0,4	5.46
98.26	4,0,0	24.57	265.88	0,8,0	14.77
99.71	0,3,0	1.45	299.12	0,9,0	33.24
122.83	5,0,0	23.12	313.89	0,0,5	14.77
125.56	0,0,2	2.73	332.35	0,10,0	18.46
132.94	0,4,0	7.39	376.67	0,0,6	44.31
147.39	6,0,0	14.45	439.44	0,0,7	62.78
166.18	0,5,0	18.79	502.22	0,0,8	62.78
171.96	7,0,0	5.78	565.00	0,0,9	62.78
188.33	0,0,3	16.38	627.78	0,0,10	62.78

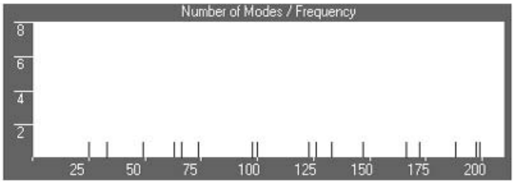


Figure 6-7. Number of modes and frequencies for a room 23 ft × 17 ft × 9 ft. (From AcousticX.)

The ratio of the dimensions determine the distribution of modes. The ratio is determined by letting the smallest dimension be 1 and dividing the other dimensions by the smallest. Obviously, the cube with its ratio of 1:1:1 results in an acoustic disaster. Even though a 12 ft (3.66 m) cube will sound differently from a 30 ft (9 m) cube, both rooms will exhibit the same modal distribution and it is the distribution that overwhelmingly determines the low frequency performance of a small room. The second room determined by the dimension of common building materials has a ratio of 1:1.5:2.0. The ratio is determined by letting the smallest dimension be 1 and dividing the other dimensions by the smallest. So an 8 ft by 16 ft by 24 ft room would have the ratio of 1:2:3. The third room, which seems reasonably good at this point, has a ratio of 1:1.89:2.56. From this we can see that in order to have a reasonable modal distribution one should avoid whole number ratios and avoid dimensions that have common factors.

6.3.1 Comparison of Modal Potency

To this point we have only considered the axial modes. The three types of modes, axial, tangential, and oblique differ in energy level. Axial modes have the greatest energy because there are the shortest distances and fewest surfaces involved. In a rectangular room tangential modes undergo reflections from four surfaces, and oblique modes six surfaces. The more reflections the greater the reflection losses. Likewise the greater the distance traveled the lower the intensity. Morse and Bolt[5] state from theoretical considerations that, for a given pressure amplitude, an axial wave has four times the energy of an oblique wave. On an energy basis this means that if we take the axial waves as 0 dB, the tangential waves are –3 dB and the oblique waves are –6 dB. This difference in modal potency will be even more apparent in rooms with significant acoustical treatment. In practice it is absolutely necessary to calculate and consider the axial modes. It is a good idea to take a look at the tangential modes because they can sometimes be a significant factor. The oblique modes are rarely potent enough in small rooms to make a significant contribution to the performance of the room.

6.3.2 Modal Bandwidth

As in other resonance phenomena, there is a finite bandwidth associated with each modal resonance. The bandwidth will, in part, determine how audible the modes are. If we take the bandwidth as that measured at the half-power points (–3 dB or $1/\sqrt{2}$), the bandwidth is[2]

$$\begin{aligned}\Delta f &= f_2 - f_1 \\ &= \frac{k_n}{\pi}\end{aligned} \tag{6-3}$$

where,

Δf is the bandwidth in hertz,

f_2 is the upper frequency at the –3 dB point,

f_1 is the lower frequency at the –3 dB point,

k_n is the damping factor determined principally by the amount of absorption in the room and by the volume of the room. The more absorbing material in the room, the greater k_n.

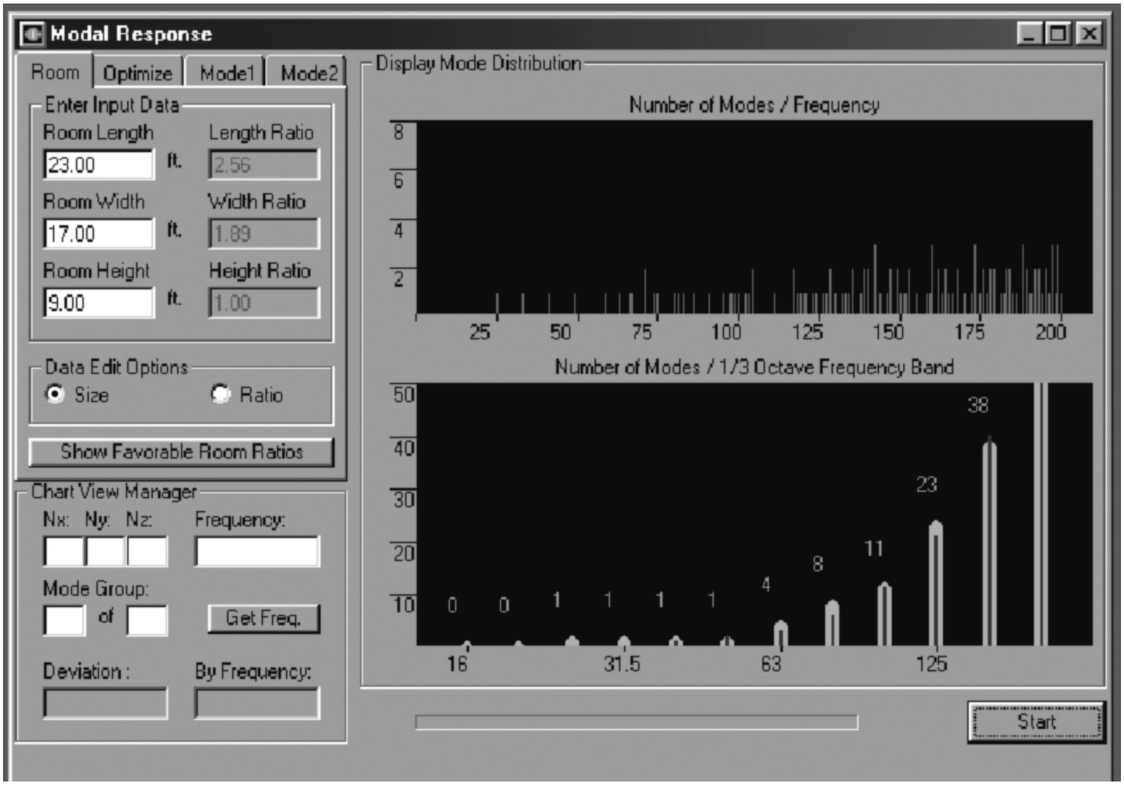
If the damping factor k_n is related to the reverberation time of a room, the expression for Δf becomes[3]

$$\begin{aligned}\Delta f &= \frac{6.91}{\pi T} \\ &= \frac{2.2}{T}\end{aligned} \tag{6-4}$$

where,

T is the decay time in seconds.

From Eq. 6-4 a few generalizations may be made. For decay times in the range of 0.3 to 0.5 s, typical of what is found in small audio rooms, the bandwidth is in the range 4.4 to 7.3 Hz. It is a reasonable assumption that most audio rooms will have modal bandwidths of the order of 5 Hz. Referring back to Table 6-6 it can be seen that in a few instances there are modes that are within 5 Hz of each other. These modes will fuse into one and occasionally some beating will be audible as the modes decay. Modal frequencies which are separated on both sides by 20 Hz or more will not fuse at all, and will be noticeable as well, although not as noticeable as a double or triple mode. Consider a room with the dimensions of 18 ft × 13 ft × 9 ft (5.48 m × 3.96 m × 2.74 m). The axial frequencies are listed in Table 6-7. There are some frequencies which double, such as 62 Hz and 125 Hz. These are obvious problems.



Modal Response

Enter Input Data

Room Length 23.00 ft. Length Ratio 2.56

Room Width 17.00 ft. Width Ratio 1.89

Room Height 9.00 ft. Height Ratio 1.00

Data Edit Options

● Size ○ Ratio

Room
Optimize
Mode1
Mode2
Show Favorable Room Ratios
Chart View Manager
Nx: Ny: Nz:
Frequency:
Mode Group:
of
Get Freq.
Deviation :
By Frequency:
Display Mode Distribution
Number of Modes / Frequency
8
6
4
2
25
50
75
100
125
150
175
200
Number of Modes / 1/3 Octave Frequency Band
50
40
30
20
10
0
0
1
1
1
1
4
8
11
23
38
16
31.5
63
125
Start

Figure 6-8. The application of Bonello's criterion 1 to the 23 ft × 17 ft × 9 ft room. From AcousticX.

However 282 Hz is also a problem frequency because it is separated by more than 20 Hz on either side.

Table 6-7. Axial Modes for a Room 18 Ft × 13 Ft × 9 Ft

Frequency	Spacing	Frequency	Spacing
31.39		219.72	2.41
43.46	12.07	251.11	31.39
62.78	19.32	251.11	0.00
62.78	0.00	260.77	9.66
86.92	24.15	282.50	21.73
94.17	7.24	304.23	21.73
125.56	31.39	313.89	9.66
125.56	0.00	313.89	0.00
130.38	4.83	345.28	31.39
156.94	26.56	347.69	2.41
173.85	16.90	376.67	28.97
188.33	14.49	376.67	0.00
188.33	0.00	391.15	14.49
217.31	28.97		

6.4 Criteria for Evaluating Room Modes

So far we have shown that there are a few general guidelines for designing small rooms with good distribution of room modes. We know that if two or more modes occupy the same frequency or are bunched up and isolated from neighbors, we are immediately warned of potential coloration problems. Over the years, a number of authors have suggested techniques for the assessment of room mores and methods for predicting the low frequency response of rooms based on the distributions of room modes. Most notably Bolt[7], Gilford,[26] Louden,[25] Bonello,[3] and D'Antonio[27] have all suggested criteria. Possibly the most widely used criteria is that suggested by Bonello.

Bonello's number one criterion is to plot the number of modes (all the modes, axial, tangential, and oblique) in ⅓ octave bands against frequency and to examine the resulting plot to see if the curve increases monotonically (i.e., if each ⅓ octave has more modes than the preceding one or, at least, an equal number). His number two criterion is to examine the modal frequencies to make sure there are no coincident modes, or, at least, if there are coincident modes, there should be five or more modes in that ⅓ octave band. By applying Bonello's method to the 23 ft × 17 ft × 9 ft room, we obtained the graph of Fig. 6-8. The conditions of both criteria are met. The monotonic increase of successive ⅓ octave bands confirms that the distribution of modes is favorable.

It is possible that the critical bands of the ear should be used instead of ⅓ octave bands. Actually, ⅙ octave bands follow critical bandwidths above 500 Hz better than do ⅓ octave bands. Bonello considered critical bands in the early stages of his work but found that one-third octave bands better show subtle effects of small changes in room dimensions.[5] Another question is whether axial, tangential, and oblique modes should be given equal status as Bonello does when their energies are, in fact, quite different. In spite of these questions, the Bonello criteria are used by many designers and a number of computer programs are using the Bonello criteria in determining the best room mode distributions.

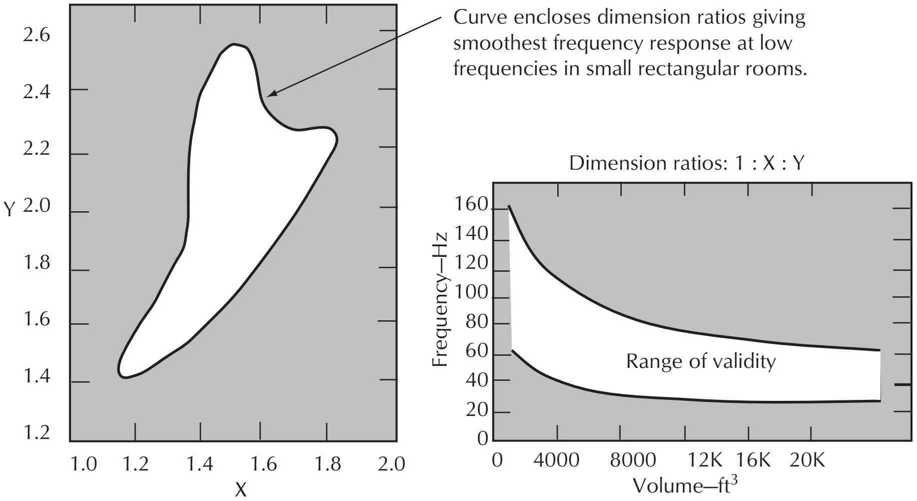
D'Antonio et al, have suggested a technique which calculates the modal response of a room, simulating placing a measurement microphone in one corner of a room then energizing the room with a flat power response loudspeaker in the opposite corner.[6] The authors claim that this approach yields significantly better results than any other criteria.

Another tool which historically has been used to help choose room dimensions is the famous Bolt footprint shown in Fig. 6-9. Please note the chart to the right of the footprint which limits the validity of the footprint. The ratios of Fig. 6-9 are all referenced to ceiling height.

6.5 Modes in Nonrectangular Rooms

Nonrectangular rooms are often built to avoid flutter echo and other unwanted artifacts. This approach is usually more expensive, therefore it is desirable to see what happens to modal patterns when room surfaces are skewed. At the higher audio frequencies, the modal density is so great that sound pressure variations throughout a rectangular room are small and there is little to be gained except, of course, the elimination of flutter echoes. At lower audio frequencies, this is not the case. The modal characteristics of rectangular rooms can be readily calculated from Eq. 6-1. To determine modal patterns of nonrectangular rooms, however, requires one of the more complex methods, such as the use of finite elements. This is beyond the scope of this book. We, therefore, refer to the work of van Nieuwland and Weber of the Philips Research Laboratories of Eindhoven, the Netherlands, on reverberation chambers.[8]

In Fig. 6-10 the results of finite element calculations are shown for 2D rectangular and nonrectangular rooms of the same area (377 ft^2 or 35 m^2). The lines are contours of equal sound pressure. The heavy lines represent the nodal lines of zero pressure of the standing wave. In Fig. 6-10 the 1,0,0 mode of the rectangular room, resonating at 34.3 Hz, is compared to a 31.6 Hz resonance of the nonrectangular room. The contours of equal pressure are decidedly nonsymmetrical in the latter. In Fig. 6-10 the 3,1,0 mode of the rectangular room (81.1 Hz) is compared to an 85.5 Hz resonance in the nonrectangular room. Increasing frequency in Fig. 6-10, the 4,0,0 mode at 98 Hz in the rectangular room is


Curve encloses dimension ratios giving smoothest frequency response at low frequencies in small rectangular rooms.

Dimension ratios: 1 : X : Y

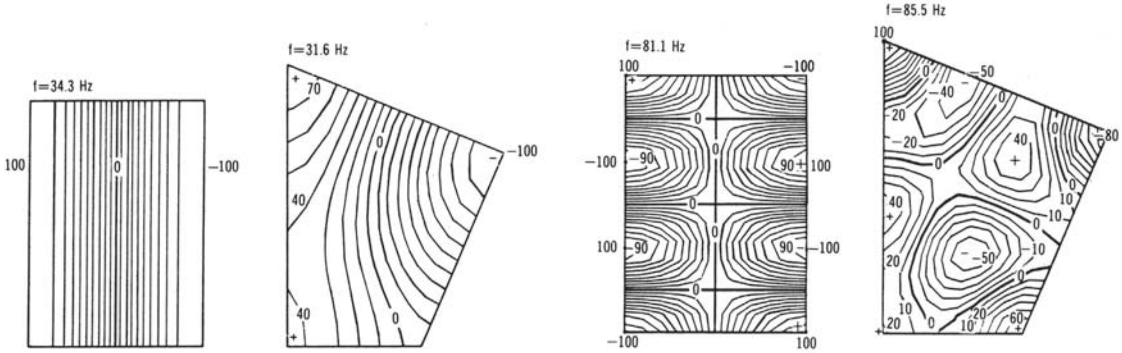
Figure 6-9. Room proportion criterion.

compared to a 95.3 Hz mode in the nonrectangular room. Fig. 6-10 shows the 0,3,0 mode at 102.9 Hz of a rectangular room contrasted to a 103.9 Hz resonance in the nonrectangular room. These pressure distribution diagrams of Fig. 6-10 give an excellent appreciation of the distortion of the sound field by extreme skewing of room surfaces.

When the shape of the room is irregular, as in Fig. 6-10, the modal pressure pattern is also irregular. The number of modes per frequency band in the irregular room is about the same as the regular room because it is determined principally by the volume of the room rather than its shape. Instead of axial, tangential, and oblique modes characteristic of the rectangular room, the resonances of the nonrectangular room all have the character of 3D (obliquelike) modes. This has been demonstrated by measuring decay rates and finding less fluctuation from mode to mode. Note that the modes did not go away and that there was not a significant change in the frequency of the modes nor in the distribution of the modes relative to frequency. What changed was the distribution of the modes in the physical space. The benefits of asymmetrical, nonrectangular designs must be measured against the drawbacks as we shall see later on in this chapter.

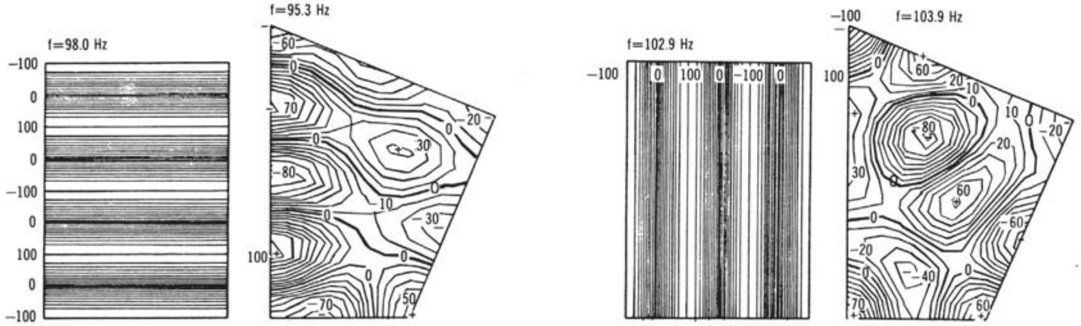
6.6 Summation of Modal Effect

Room modes determine the performance of small rooms below f_l. The following criteria should be applied when evaluating room ratios or dimensions in terms of modal distribution. When considering axial modes, there should be no modes within 5 Hz of each other, and no mode should be greater than 20 Hz from another. Since the modal bandwidth in small rooms is approximately 5 Hz, any modes that are within 5 Hz of each other will effectively merge into one. Modes that are isolated by more than 20 Hz will not have masking from any other modes nearby and will likely stand out. Obviously there should not be any double or triple modes. Some criteria


f=34.3 Hz
100
0
−100
f=31.6 Hz
70
0
−100
40
40
0
f=81.1 Hz
100
−100
0
−100
−90
90
100
100
−90
90
−100
−100
100
f=85.5 Hz
100
0
−50
−40
20
−20
−80
40
0
10
−10
40
−50
−10
20
0
10
20
0
10
20
60

A. The 1,0,0 mode of the rectangular room (34.3 Hz) compared to the nonrectangular room (31.6 Hz).

B. The 3,1,0 mode of the rectangular room (81.1 Hz) compared to the nonrectangular room (85.5 Hz).


f=98.0 Hz
−100
0
100
0
−100
0
100
0
−100
f=95.3 Hz
60
0
70
−20
0
30
−80
10
−30
100
0
−70
50
f=102.9 Hz
−100
0
100
0
−100
0
−100
f=103.9 Hz
100
60
20
10
0
10
−20
80
−20
30
60
−60
−20
−40
0
70
60

C. The 4,0,0 mode in the rectangular room (98 Hz) compared to the nonrectangular room (95.3 Hz).

D. The 0,3,0 mode (102.9 Hz) contrasted to the nonrectangular room (103.9 Hz).

Figure 6-10. A comparison of calculated 2D sound fields in rectangular and nonrectangular rooms having the same areas. After Reference 8.

should also be applied to all the modes, the axial, tangential, and oblique. There are many excellent tools for calculating modal distribution.

6.7 The Question of Reverberation

W.C. Sabine, who first formulated the equation to calculate reverberation time, described reverberation in this way: "reverberation results in a mass of sound filling the whole room and incapable of analysis into its distinct reflections."[10] What Sabine was saying, although he did not use these terms, was that for true reverberation to exist, there needs to be a homogenous and isotropic sound field. Usually such conditions are approached in physically large rooms that do not contain much absorption. Unfortunately the term reverberation is popularly understood to be equivalent to decay. Does reverberation time refer to the decay of a well-established, totally homogenous, diffuse sound field that exhibits no net energy flow due to the richness of the reflections present or does reverberation time refer to the decay of any sound in a room no matter what the nature of the sound is, even if it is not diffuse? To some extent, this is a question of semantics. It is interesting to note that maybe Sabine himself perhaps anticipated the confusion that would eventually arise because in the same paper he wrote:

> *The word "resonance" has been used loosely as synonymous with "reverberation" and even with "echo" and is so given in some of the more voluminous but less exact popular dictionaries. In scientific literature the term has received a very definite and precise application to the phenomenon where ever it may occur. A word having this significance is necessary; and it is very desirable that the term should not, even popularly, by meaning many things, cease to mean anything exactly.*[11]

It is the opinion of this author that this is precisely where we find ourselves today. Without a rigorous definition and application of the concept of reverberation, we are left with something which ceases to mean anything exactly.

When Sabine first measured the decay of the reverberation in Fogg Lecture Hall at Harvard, he did it with an organ pipe and a stopwatch. He had no way of examining the fine detail of the reflections or any of the components of the sound field, nor was he initially looking at decay as a function of frequency. (Later on he looked at decay as a function of frequency, but never connected this to room size or shape.) He could only measure the decay rate of the 513 Hz pipe he was using. The volume of the lecture hall was approximately 96,700 ft^3.[13] The room was large enough that 512 Hz was not going to energize any of the normal room modes. Since there was virtually no absorption in the room whatsoever, it is likely that Sabine was measuring a truly diffuse sound field. It is interesting to note that in Sabine's early papers he rarely mentions the dimensions other than the volume of the rooms he was working in. He was convinced that it was the volume of the room that was important. The mean free path was also central to his thesis. The *MFP* is defined as the average distance a given sound wave travels in a room between reflections.[14] The equation for finding the mean free path is

$$MFP = \frac{4V}{S} \tag{6-5}$$

where,
V is the volume of the room,
S is the total surface area.

Consider a small room with dimensions of 12 ft × 16 ft × 8 ft high (3.66 m × 4.88 m × 2.44 m). This room will have a volume of 1536 ft^3 (43.5 m^3) and a total surface area of 832 ft^2 (77.3 m^2). Putting these numbers into Eq. 6-5 yields a result of a MFP of about 7.38 ft. At the average speed of sound (1130 ft/s or 344 m/s) this distance will be covered in 0.00553 s or 5.53 ms. It is generally accepted that in small rooms, after approximately four to six bounces, a sound wave will have lost most of its energy to the reflecting surfaces and will become so diffuse as to be indistinguishable from the noise floor. This of course depends on the amount of absorption in the room. In very absorptive rooms there may not be even two bounces. In very live hard rooms a wave may bounce more than six times. In this room a single wave will take only 32.6 ms to bounce five times and be gone. Compare this with a large room. Consider a room that is 200 ft long by 150 ft wide with a 40 ft ceiling (61 m × 45.7 m 12.2 m). This room will have a *MFP* of 54.5 ft (16.61 m). It will take 241.3 ms for a single wave to bounce five times and dissipate.

Sabine was not interested in the shape of the room or even in the distribution of the absorptive material. He focused on the statistical nature of the diffuse sound field and on the rate of decay. Other researchers looked at similar issues eventually dividing the time domain performance into smaller and smaller regions and examining their contributions to the subjective performance of rooms.

Fig. 6-11 is an Envelope* Time Curve of a large, reverberant room, measured using time delay spectrometry. The left side of the graph represents *t*0 or the beginning of the measurement. Note that this is the point in time what the signal leaves the loudspeaker. If the microphone represents the observer in this system, the observer would not hear anything until *t*(0) + *t*(*x*) where *t*(*x*) is the time of takes for the sound to leave the loudspeaker and arrive at the observation point. In this measurement, that time is 50 ms because the loudspeaker was about 56 ft (17 m) away from the microphone. This first arrival is known as the direct sound because it is the sound energy that first arrives at the listener or microphone, before it reflects off of any surface. A careful examination of this graph shows a small gap between the direct sound and the rest of the energy arriving at the microphone. This is known as the initial time gap (ITG) and it is a good indicator of the size of the room. In this room it took about 50 ms for the sound to travel from the loudspeaker to the microphone then another 40 ms (90 ms total) for sound to leave the speaker and bounce off of some surface to arrive at the microphone. Therefore, in this room the ITG is about 40 ms wide.

Fig. 6-12 is an enlargement of the first 500 ms of Fig. 6-11. The ITG can be clearly seen and is about 40 ms long. The sound then takes about 130 ms or so to build up to a maximum at around 270 ms. Fig. 6-12 shows that the sound then decays at a fairly even rate over the next 4 s till the level falls into the noise floor.

If we perform a Schroeder integration[28] of the energy then measure the slope and extrapolate down to 60 dB below the peak, we see the reverb time of this room to be on the order of 6.8 s at 500 Hz, Fig. 6-13.

It is useful to take a look at the ETC of a small room for comparison, Fig. 6-14.

Careful examination of Fig. 6-14 reveals a room dominated by strong discrete reflections that start coming back to the observation point within a few ms of the direct sound. By 30 ms after the direct sound, the energy has decayed into the noise floor.

Since there is no significant diffuse or reverberant field in acoustically small rooms, equations having reverb time as a variable are not appropriate.

It is important to understand that small rooms must be treated differently with respect to frequency.

Consider Fig. 6-15. These boundaries should not be understood as absolute or abrupt. They are meant to serve as guidelines and the transitions from one region to another are actually very gradual.

Region 1 is the region from 0 Hz up to the first mode associated with the longest dimension. In this region there is no support from the room at all, and there is not much one can do to treat the room. Region 2 is bounded

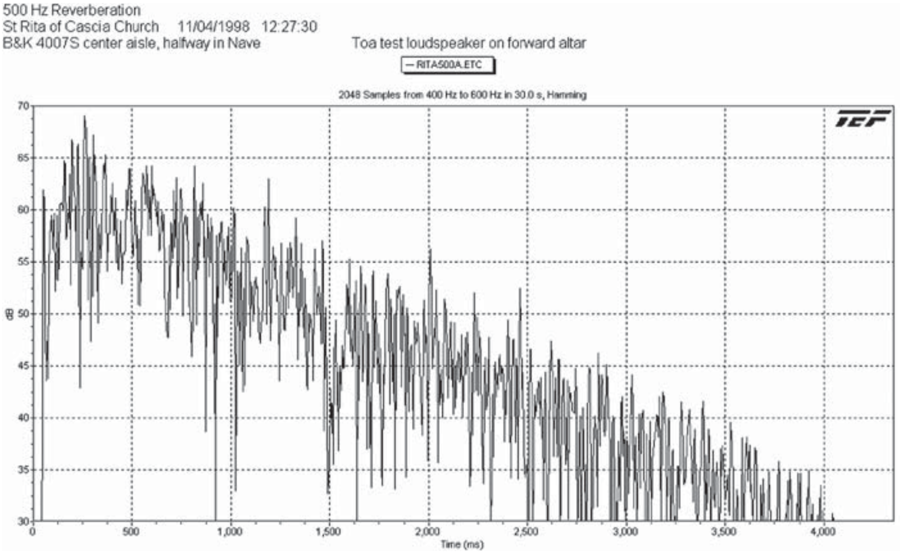


Figure 6-11. ETC of a large reverberant church. Measurement courtesy of Jim Brown.

---

* The term Energy Time Curve was suggested by Heyser and adopted by Crown and Gold Line, respectively, in their Time Delay Spectrometry software. Recently Davis and Patronis have suggested that Envelope Time Curve is a better label for this graph.

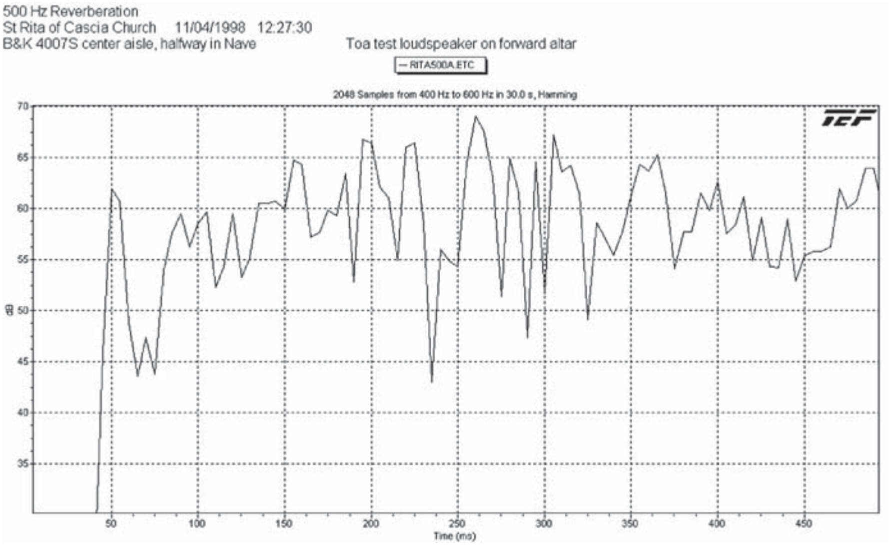


Figure 6-12. Enlargement of the first 500 ms of Fig. 6-11.

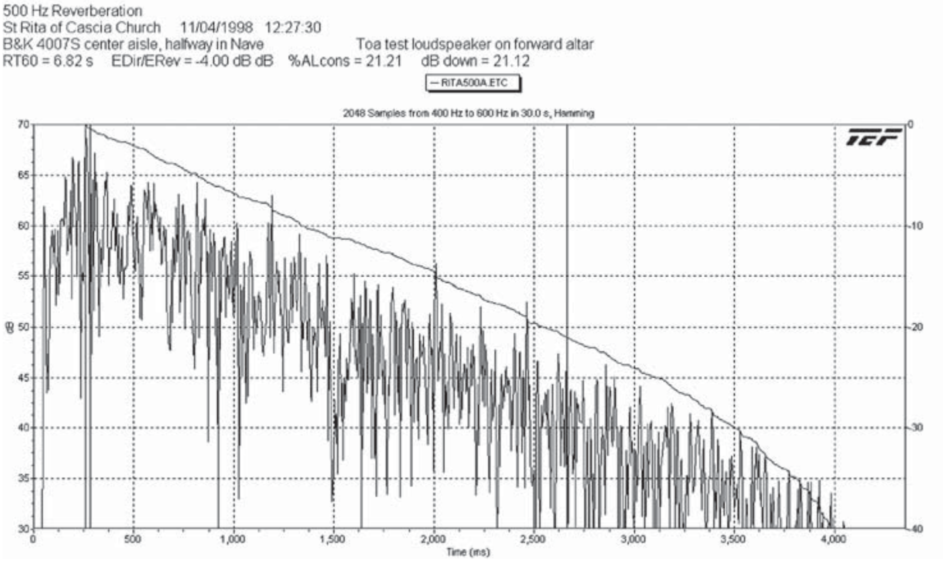


Figure 6-13. Schroeder integration of Fig. 6-11.

by the first mode on the low end of the spectrum and *f* at the high end, where $f \approx 3C/RSD$ (rooms smallest dimension). In this region where the room modes dominate the acoustical performance, wave behavior is the best model and some forms of bass absorption can work well. Region 3 spans from *f* to roughly four times *f*. This region is dominated by diffraction and diffusion. The final region is where the wavelengths are generally small relative to the dimensions of the room. In this region, one can use a ray acoustics approach as we are dealing with specular reflections.

The discussion of how to quantify the decay of sound in small rooms continues. Most recording studios, control rooms, and listening rooms are too small to have a completely diffuse sound field, especially at the lower frequencies. In small, acoustically dead rooms the only frequencies that have any significant decay are those that are at or very near to the natural resonances of the room. This decay time is,

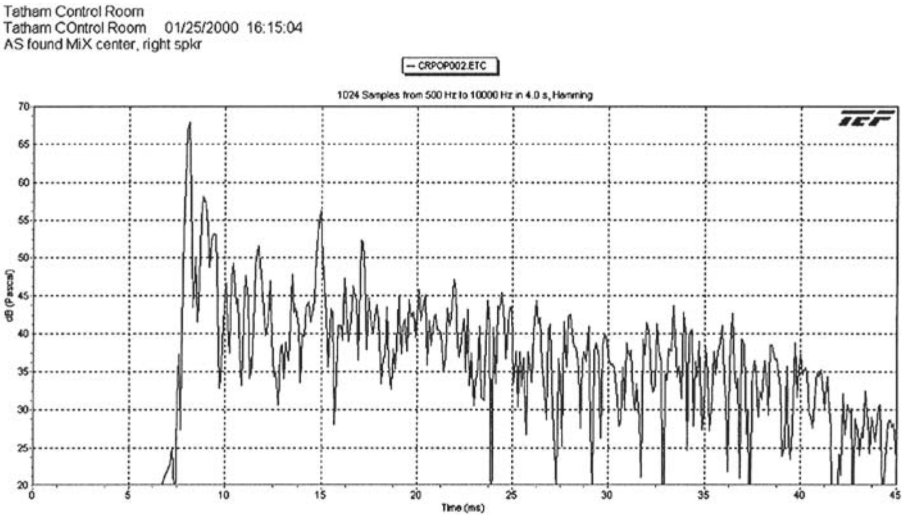


Figure 6-14. ETC of a small room, approximately 250 ft² area.

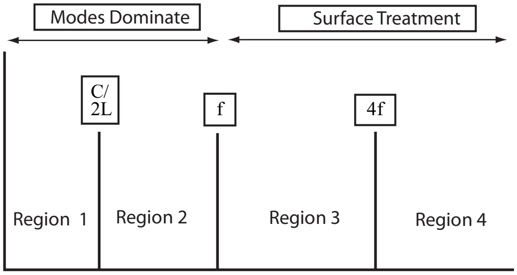


Figure 6-15. Regions for room treatment.

strictly speaking, reverberation and should be treated as such. The Sabine equation and its offspring are not going to help in predicting how much absorption is going to be needed to modify a room. More discussion and research are needed to be able to fully quantify the behavior of absorption in small rooms.

As is mentioned in Chapter 5, the standard way to measure acoustical treatment is according to ASTM C 423-00. This is an indirect method that looks at the impact that the material has on a diffuse sound field. In rooms that do not exhibit any diffuse sound field in the frequency range of interest, we need another way to measure acoustic treatment. An alternative method is outlined in Chapter 5.

6.7.1 *Room Shape*

We have referred to the statistical approach (e.g., reverberation) and the wave approach (modes) to acoustical problems, and now we come to the geometrical approach. The one overriding assumption in the application of geometrical acoustics is that the dimensions of the room are large compared to the wavelength of sound being considered. In other words, we must be in region 3 of Fig. 6-15 where specular reflection prevails and the short wavelengths of sound allow treating sound like rays.

A sound ray is nothing more than a small section of a spherical wave originating at a particular point. It has a definite direction and acts as a ray of light following the general principles of geometric optics. Geometrical acoustics is based on the reflection of sound rays. This is where the shape of the room is the controlling acoustical aspect. Like the quest for room ratios the search for the perfect room shape is also ellusive. Some have suggested that nonparallel surfaces are a must, however there are no perfect shapes. There are some shapes that work well for some applications.

6.7.2 *Reflection Manipulation*

In open space (air-filled space, of course) sound travels unimpeded in all directions from a point source. In real rooms we don’t have point sources, we have loudspeakers or other sound sources such as musical instruments that do not behave like the theoretical point source. Real sources have characteristic radiation or directional patterns. Of course in real rooms the sound does not travel unimpeded for very long, depending on the MFP. After the sound leaves its source it will bounce off of some surface and will interact with the unreflected sound. This interaction can have a profound impact on the perception of the original sound. There is an elegant way to model the reflections in a room. The reflection can be considered to come from an image of the source on the opposite side of the reflecting surface and equidistant from it. This is the simple case: one source, one surface, and one image. If this reflecting surface is now taken to be one wall of a room, the picture is immediately complicated. The source now has an image in each of five other surfaces, a total of six images sending energy back to the receiver. Not only that, images of the images exist and have their effect, and so on. A physicist setting out to derive the mathematical expression for sound intensity from the source in the room at a given receiving point in the room must consider the contributions from images, images of the images, images of the images of the images, and so on. This is known as the image model of determining the path of reflections. The technique is fully described in Chapter 9.

6.7.3 Comb Filters

When the direct sound and a reflection combine at some observation point, a spectral perturbation often called a comb filter is produced. The frequency of the first notch and the spacing of the rest of the notches is base on the delay between the two arrivals. The first notch *F* in hertz is calculated by

$$F = \frac{1}{2t} \tag{6-6}$$

where,
t is the delay in seconds.

Each successive notch will be at

$$\frac{1}{t}$$

Fig. 6-16 shows the response of a system with a delay of 1.66 ms between the two signals. Reflections can dramatically change the way program material sounds depending on the time of arrival, the intensity, and the angle of incidence relative to the listener. For a more in-depth treatment of how comb filters are created, the reader is referred to reference 16.

In 1971 M. Barron wrote a paper exploring the effects of reflections on the perception of a sound.[17] He was trying to quantify the effects of lateral reflections in concert halls. Although his work was conducted in large reverberant spaces, a number of small room designers look to his work with great interest as he is considering the first 100 ms of a sound field in a room. In small rooms that is often all you will get. Fig. 6-17 is a graphic summary of the effects of a single lateral reflection. It can be seen that the very early reflections, on the order of 0 to 5 ms, can cause image shifts even when very low in amplitude relative to the direct sound. This can be important as one considers, for example, the accepted practice of placing loudspeakers on the meter bridge of a recording console.

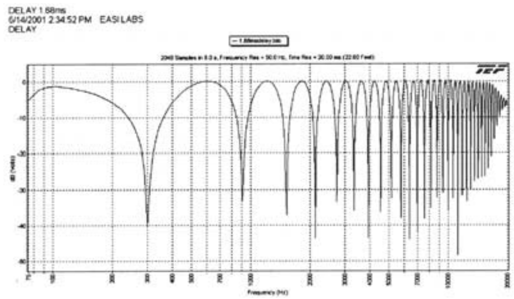


Figure 6-16. Response of a system with a delay of 1.66 ms between the two sources.

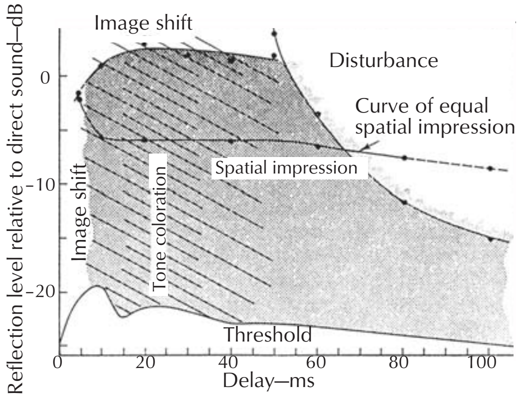
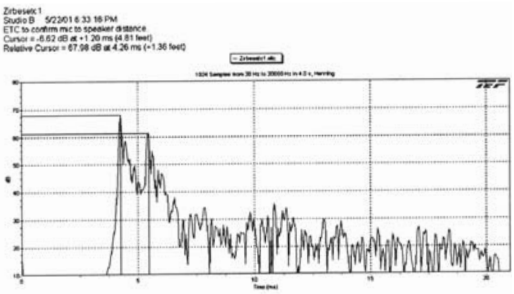
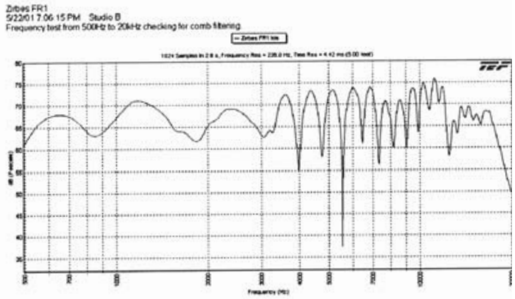


Figure 6-17. Graphic summary of the effects of a single lateral reflection.

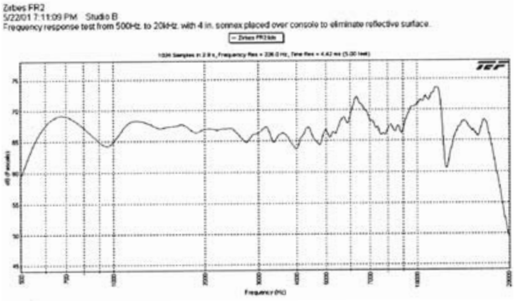
Fig. 6-18A shows an ETC of a popular nearfield loudspeaker placed on the meter bridge of a recording console, measured at the mix position. The first spike is of course the direct arrival of the signal from the loudspeaker. The second spike is the reflection off of the face of the console, approximately 1.2 ms later. Fig. 6-18B shows the resulting frequency response when these two signals arrive at the microphone. Finally C is the frequency response of three loudspeaker with the reflection removed. This author finds it curious indeed that this practice of placing a loudspeaker on the console ostensibly to remove the effects of the room and get a more accurate presentation actually results in seriously coloring the response of the speaker, and will



A. ETC of the "near field" loudspeaker on a meter bridge.



B. Frequency response of the "near field" loudspeaker on a meter bridge.



C. Frequency response of the "near field" loudspeaker on the meter bridge with the console covered with acoustical absorbent, removing the reflection.

Figure 6-18. The effect of reflections bouncing off a studio console. Measurement courtesy of Mathew Zirbes.

have a significant impact on the ability to accurately perceive a stereo image.

Although Barron did not look at the effect of reflections arriving from below as in the case of a console reflection, the effect is clearly audible. In 1981 C.A.P. Rodgers[18] noted a similarity between the spectral notches created as a result of loudspeaker misalignment and those created by the pinna which have been shown to play an important role in localization. She postulated that the presence of spectral notches would impair or at least confuse the auditory system's ability to decode position cues. This could explain the phenomenon noted by Barron. The very early reflections are those that cause notches similar in spectral positioning to those caused by the pinna. These are the reflections that cause image shifts.

6.8 Small Room Design Factors

The basic tools for looking at small room performance have been addressed. We now turn our attention to small room design factors. We have divided small rooms into three broad categories; precision listening rooms, rooms for microphones, and rooms for entertainment. We are not trying to imply that there is only one way to build a control room or an entertainment room. There are different design criteria for different outcomes. The categories presented here are not intended to be exhaustive; rather they are intended to be general and representative. It should also be noted that we are not including the noise control issues that are often an important part of room design. The reader is referred to Chapter 3 for noise control information.

6.8.1 The Control Room Controversy

Since most control rooms are acoustically small, it is appropriate to discuss control room design in general in this context. Some insist that control rooms should be as accurate as possible. Others insist that since music is rarely listened to in highly precise analytic rooms, recorded music would be better served if control rooms were more like entertainment rooms; not so sterile, but rather designed so that everything sounds subjectively great. Indeed many recordings are made in rooms that are not close to precision listening rooms. This debate will probably never be resolved as long as there are deductive and inductive reasoners, left-brain and right-brain people, artists and engineers. In the next few sections we are not attempting to solve this debate, rather we are trying to set out some simple guidelines. The most important task for the room designer is to listen to the client and not make assumptions about what it is he or she is looking for.

6.8.2 Precision Listening Rooms

These are rooms where the primary goal is for the listener to have as much confidence as possible that what is heard is precisely what is being or has been recorded. Frequently, users of these rooms are performing tasks that require listening analytically to the program and making decisions or judgments about what is heard.

Examples of rooms in this category are recording control rooms, mastering suites, and audio production rooms. The state of the art at this writing does not permit us to design transducers or electronics that are perfect so as to afford the user 100% confidence that what is heard is fully equivalent to what has been or is being recorded. We can, however, design rooms that fully meet this criterion. An anechoic chamber would indeed be 100% neutral to the loudspeaker, allowing the user to hear precisely and only what is coming out of the speaker. The problem is that anechoic chambers are quite possibly the most acoustically hostile places we can imagine. It is difficult to spend a few minutes in an anechoic chamber let alone try to be creative and make artistic decisions about music in one. The challenge is to build a room that will not significantly interact with the loudspeaker by means of room modes or reflections that arrive at the listening position and still be a place that is subjectively acceptable to the user. There have been a number of good approaches to this problem over the years starting with LEDE™,[19] Reflection Free Zone™ (RFZ),[20] and the Selectively Anechoic Space™.[21] Later came Tom Hidley's neutral room or nonenvironment design and more recently, David Moulton has proposed his wide-dispersion design. These approaches all endorse attenuating or completely eliminating all the early reflections, creating a space that is essentially anechoic when energized by the loudspeakers and listened to in the prescribed position, but in all other ways it is an average room. Reflections can be eliminated or reduced at the listening position by changing the angle of the reflector, by using an absorber, or by using a diffuser. It should be noted that Angus questioned the use of diffusion in controlling lateral reflections.[22]

On the surface one might wonder why all sound rooms are not built this way. The reason is that most people do not listen to music analytically. In precision rooms, music that is poorly recorded will sound that way. One can certainly design rooms where the music sounds better than it does in a precision room. There are artifacts that one can build into a room that are subjectively very pleasing, but they are part of the *room* and not the recording. The recording engineer generally wants to know what exactly is in the recording. The engineer generally listens to the product in a number of different environments before releasing it to insure that it does hold up even under nonideal conditions.

So-called good sounding artifacts can be observed in the frequency domain as well as the time domain. For example, if a room has an audible room mode at 120 Hz music might sound full and rich in the upper low end and be quite pleasing, however the fullness is in the room, not the recording. The recording may in fact be "thin" or lacking in the low end because the room is adding to the mix. In the time domain, a reflection that occurs in the first 10 ms or so and comes from the side (a lateral reflection) might result in a perception of a stereo image that is much wider than the physical separation of the speakers might allow. This might be perceived as a very good sound stage, but it is an artifact of the room and not of the recording.[23]

Designing such a room is an art and a science. It is beyond the scope of this book to detail a complete room design protocol, however, the steps in designing such a room must include:

1. Choosing a set of room ratios that yield a modal distribution that will result in the best possible low frequency performance.
2. Choosing a symmetrical room shape so that each loudspeaker interacts with the room in exactly the same way.
3. Choosing and placing acoustical treatment so that the early reflections (at least the first 18 ms) are attenuated at and are at least 18 dB below the direct sound. Care should be taken to insure that the treatment chosen exhibits a flat absorption characteristic at the frequency range of interest and at the angles of incidence. The energy time curve should be measured to insure that the direct sound is not compromised over the entire listening area.
4. Placing equipment and furniture in the room in such a way as to not interfere with the direct sound. It should be noted that the recording console is often the most significant acoustical element in the control room.
5. Insuring that there are enough live and diffusive surfaces in the room so that the overall subjective feel of the room is that of a normal room and not an anechoic chamber.

6.8.3 Rooms for Microphones

Designers are frequently asked to design rooms that are intended for recording or use with live microphones. Recording studios, vocal booths, and even conference rooms could be part of this category. The criteria in these rooms are almost all subjective. End users want rooms that sound good and that are comfortable to work in. The acoustician is well advised to work with a good interior decorator as a significant part of what makes someone feel comfortable in a room is the way the room is decorated and lit. Obviously noise control is a large part of the design criteria. There are a few general rules that will help with the acoustics of these small rooms:

1. Like the precision rooms, these rooms will work better if the proportions of the room result in optimal modal distribution.
2. Unlike the precision rooms, studios and vocal booths often work best when they are *not* symmetrical.
3. Avoid parallel surfaces if possible.
4. Use treatment that is as linear as possible, both statistically and by direct measurement of reflected sound.
5. Avoid treating entire surfaces with a single form of treatment. For example, covering an entire wall with an absorber will usually be less effective than treating some areas and leaving some alone.
6. Listen carefully to the kinds of words the end user employs to describe the space either in terms of what is desired or in terms of something that need modification. Words like *intimate*, *close*, *dark*, *dead*, *quiet* are usually associated with the use of absorption. Words like *open*, *live*, *bright*, *airy* are often used in conjunction with diffusion.
7. Placing absorption in the same plane as the microphone will increase the apparent MFP and result in a longer ITG (initial time gap). This often makes the room seem larger. For example, in a vocal booth that is normally used by standing talent, place the absorption on the walls such that both the talent and the microphone are in the same plane as the absorptive area. In a conference room placing a band of absorption around the room at seated head height will help improve the ability to communicate in the room.

6.9 Rooms for Entertainment

There was a serious temptation to call this section "Rooms That Sound Good." The temptation was resisted to avoid the criticism that the section titles would thus imply that precision rooms don't sound good. It is a matter of goals. As was pointed out, the purpose of the precision room is analysis. This section will cover rooms that are designed for entertainment.

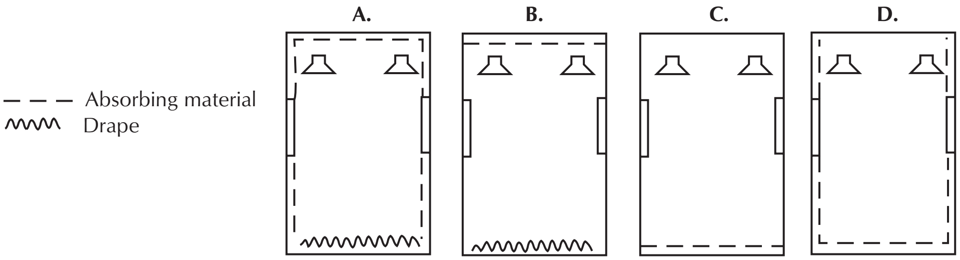
Of course it is much more difficult to set out design criteria for a good sounding room. As with any subjective goal it comes down to the tastes and preferences of the end user. To a great extent how one approaches an entertainment room depends on the type of system to be used, and the type of entertainment envisioned. An audiophile listening room will be treated differently from a home theater. It should be noted that in the world of home entertainment there exists a very rich audio vocabulary. Some of the words that are used like *spaciousness* and *localization* have meanings that are consistent with the use of these words in the scientific audio community. Subjective words like *air*, *grain*, *definition*, *impact*, and *brittleness* are much more ambiguous and are not yet mapped into the physical domain so that we know how to control them. One of the challenges is when the end user wants two mutually exclusive aspects optimized! The so-called Nippon-Gakki experiments of 1979[24] quite elegantly showed how different subjective effects can be created by simply moving acoustic treatment to different locations in a room, Fig. 6-19. Note that when localization is rated good, spaciousness is rated poor and vice versa.

Some general points:

1. In home entertainment systems the distribution of room modes is somewhat less important. Having modal support in the low end although inaccurate can result in rooms that sound fuller. This might enhance a home theater system.
2. Absorption should be used sparingly. These rooms should be quiet, not dead. If absorption is to be used, it must be linear.
3. Remember that everything in a room contributes to the acoustics of the room. Most home entertainment rooms will have plush furniture that will be a significant source of absorption. The furnishings should be in place before the final treatment is considered.

The furnishings should be in place before the final treatment is considered.

1. Lateral reflections should be emphasized by using critically placed diffusers. Lateral reflections can dramatically increase the sense of spaciousness in a room.
2. Absorptive ceilings tend to create a sense of intimacy and a feeling of being in a small space. If this is not desired, use some absorption to control the very early reflections but leave the rest live.



	A.	B.	C.	D.
Localization	Good	Poor	Poor	Fair
Noncoloration	Good	Poor	Poor	Fair
Loudness	Poor	Good	Good	Good
Broadening of image	Poor	Good	Good	Fair
Perspective	Good	Good	Poor	Good

Figure 6-19. Summary of the results of the Nippon-Gakki psychoacoustical experiments.

References

1. L. L. Beranek, *Acoustics*, Second Edition, New York: McGraw-Hill, 1986.
2. Ibid. p. 292.
3. O. J. Bonello, "A New Criterion for the Distribution of Normal Room Modes," *Journal of the Audio Engineering Society*, vol. 29, no. 9, pp. 597–606, September 1981.
4. Ibid.
5. Ibid.
6. F. A. Everest, *Master Handbook of Acoustics,* Third Edition, TAB Books, 1998.
7. R. H. Bolt, "Note on Normal Frequency Statistics for Rectangular Rooms," *JASA*, vol. 18 no. 1, July 1946, pp. 130-133.
8. J. M. van Nieuwland and C. Weber, "Eigenmodes in Nonrectangular Reverberation Rooms," *Noise Control Engineering*, vol. 13, no. 3, pp. 112 – 121, November-December 1979.
9. AcousticX available from Pilchner/Schoustal, Toronto, Canada.
10. W. C. Sabine, *Collected Papers on Acoustics*, Peninsula Publishing, Los Altos, CA, 1992, pp. 9.
11. Ibid. pp. 8 and 9.
12. Oxford English Dictionary.
13. W. C. Sabine, *Collected Papers on Acoustics,* Peninsula Publishing, Los Altos, CA, 1992.
14. L. L. Beranek, *Acoustics*, Second Edition, New York: McGraw-Hill, 1986.
15. M. R. Schroeder "Frequency – Correlation Functions of the Frequency Responses in Rooms" *Journal of the Acoustical Society of America,* vol. 34/12 (1962).
16. F. A. Everest, *Master Handbook of Acoustics,* Third Edition, TAB Books, 1998.
17. M. Barron, "The Subjective Effects of First Reflections in Concert Halls- The Need for Lateral Reflections," *Journal of Sound and Vibration*, vol. 15/4 pp. 475-494 (1971).
18. C. A. P. Rodgers "Pinna Transformations and Sound Reproduction," *Journal of the Audio Engineering Society*, Vol 29 pp. 226-234 (1981).
19. Don Davis and Chips Davis, "The LEDE Concept for the Control of Acoustic and Psychoacoustic Parameters in Recording Studio Control Rooms." *Journal of the AES*, vol. 28, no. 9, pp. 585, 595. September 1980
20. "The RFZ/ RPG Approach to Control Room Monitoring." 76 AES Convention, NY 1984, preprint 2157.
21. Jones, Kendall, and Martins, "Designing a Room for Spatial Hearing Research using the TEF," Proceedings of 1984 International Computer Music Conference, IRCAM, Paris.
22. J. A. Angus "The Effects of Specular vs. Diffuse Reflections on the Frequency Response at the Listener," *Journal of the Audio Engineering Society,* vol. 49, no. 3, March 2001.
23. Jones, Kendall, and Martins, "Optimizing Control Rooms for Stereo Imagery," Paper presented at 1985 Acoustical Society Convention, Nashville TN.

24. Kishinaga, Shimizu, Ando, and Yamaguchi, "On the Room Acoustic Design of Listening Rooms," Audio Engineering Society 64th Convention (Nov. 1979) preprint #1524.
25. M. M. Louden, "Dimension Ratios of Rectangular Rooms with Good Distribution of Eigentones," *Acoustica*, no. 24, pp. 101–104, 1971.
26. C. L. S. Gilford, "The Acoustic Design of Talk Studios and Listening Rooms," *Journal of the Audio Engineering Society*, vol. 27, pp. 17–31, 1979.
27. Cox and D'Antonio, "Room Sizing and Optimization at Low Frequencies," *Journal of the Acoustical Society of America,* vol. 52, no. 6, pp. 640–653, June 2004.
28. M. R. Schroeder, "New Method of Measuring Reverberation Time," *Journal of the Audio Engineering Society*, vol. 35, pp. 299–305, 1987.

Chapter 7

Acoustics for Auditoriums and Concert Halls

by Dr. Wolfgang Ahnert and Hans-Peter Tennhardt

7.1 Introduction . . . 147
7.2 Room-Acoustical Criteria, Requirements . . . 147
  7.2.1 Time Criteria . . . 148
    7.2.1.1 Reverberation Time RT_{60} . . . 148
    7.2.1.2 Bass Ratio (BR) (Beranek) . . . 151
  7.2.2 Energy Criteria . . . 152
    7.2.2.1 Strength Measure (G) (P. Lehmann) . . . 153
    7.2.2.2 Sound Pressure Distribution (ΔL) . . . 153
    7.2.2.3 Interaural Cross-Correlation Coefficient (IACC) . . . 153
    7.2.2.4 Center Time (ts) (Kürer) . . . 154
    7.2.2.5 Echo Criterion (EK) (Dietsch) . . . 154
    7.2.2.6 Definition Measure C_{50} for Speech (Ahnert) . . . 154
    7.2.2.7 Speech Transmission Index (STI) (Houtgast, Steeneken) . . . 155
    7.2.2.8 Articulation Loss, Alcons, with Speech (Peutz, Klein) . . . 156
    7.2.2.9 Subjective Intelligibility Tests . . . 156
    7.2.2.10 Clarity Measure (C80) for Music (Abdel Alim) . . . 159
    7.2.2.11 Sound Coloration Measures (KT) and (KH) for Music (Schmidt) . . . 160
    7.2.2.12 Spatial Impression Measure (R) for Music (U. Lehmann) . . . 160
    7.2.2.13 Lateral Efficiency (LE) for Music (Jordan), (LF) (Barron) and (LFC) (Kleiner) . . . 161
    7.2.2.14 Reverberance Measure (H) (Beranek) . . . 161
    7.2.2.15 Register Balance Measure (BR) (Tennhardt) . . . 162
7.3 Planning Fundamentals . . . 162
  7.3.1 Introduction . . . 162
  7.3.2 Structuring the Room Acoustic Planning Work . . . 164
    7.3.2.1 General Structure . . . 164
    7.3.2.2 Room Form and Sound Form . . . 165
  7.3.3 Primary Structure of Rooms . . . 165
    7.3.3.1 Volume of the Room . . . 165
    7.3.3.2 Room Shape . . . 167
7.3.3.2.1 Ground Plan . . . 168
7.3.3.2.2 Ceiling . . . 171
7.3.3.2.3 Balconies, Galleries, Circles . . . 173
    7.3.3.3 Room Topography . . . 173
7.3.3.3.1 Sloping of Tiers, Sight Lines . . . 173
7.3.3.3.2 Platform Configuration in Concert Halls . . . 175
7.3.3.3.3 Orchestra Pit . . . 177

7.3.4 Secondary Structure of Rooms 178
  7.3.4.1 Sound Reflections at Smooth Plane Surfaces 178
  7.3.4.2 Sound Reflection at Smooth Curved Surfaces 179
  7.3.4.3 Sound Reflections at Uneven Surfaces 180
  7.3.4.4 Sound Absorbers 183
7.3.4.4.1 Sound Absorption Through Porous Materials 184
7.3.4.4.2 Sound Absorption by Panel Resonances 184
7.3.4.4.3 Helmholtz Resonators 185
7.3.4.4.4 Sound Absorption by the Audience 186
7.4 Variation of Room Acoustics by Construction or Electroacoustic Methods 188
7.4.1 Variable Acoustics 188
  7.4.1.1 Fundamentals of Variable Acoustics 188
7.4.2 Electronic Architecture 190
  7.4.2.1 Use of Sound Delay Systems for Enhancing Spaciousness 191
7.4.2.1.1 Ambiophony 191
7.4.2.1.2 ERES (Electronic Reflected Energy System) 191
  7.4.2.2 Travel-Time-Based Reverberation-Time Enhancing Systems 191
7.4.2.2.1 Assisted Resonance 192
7.4.2.2.2 Multi-Channel-Reverberation, MCR 192
  7.4.2.3 Modern Procedures for Enhancing Reverberation and Spaciousness 193
7.4.2.3.1 Acoustic Control System (ACS) 193
7.4.2.3.2 Reverberation on Demand System, RODS 193
7.4.2.3.3 LARES 194
7.4.2.3.4 System for Improved Acoustic Performance (SIAP) 194
7.4.2.3.5 Active Field Control, AFC 194
7.4.2.3.6 Virtual Room Acoustic System Constellation™ 195
7.4.2.3.7 CARMEN® 196
7.4.3 Conclusions and Outlook 197
References 197

7.1 Introduction

With musical or spoken performances in auditoriums and concert halls, acoustic evaluation is mainly based on the subjective perception of audience and performers. These judgements are generally not based on defined criteria, but characterize the sensed tonal perception. Besides the secondary factors influencing the overall acoustic impression like, for instance, comfortableness of the seats; air conditioning; interference level; and optical, architectural, and stylistic impression, it is especially the expectation of the listener that plays a significant role for the acoustic evaluation. If a listener in a classical concert is sitting next to the woodwind instruments but hears the brass instruments much louder, even though he cannot see them, his expectations as a listener, and thus the acoustics are off. Numerous subjective and objective room-acoustical criteria were defined and their correlation determined in order to objectify these judgments. However, these individual criteria are closely linked with each other and their acoustic effects can neither be exchanged nor individually altered. They become effective for judgment only in their weighted totality. The judgment of the performers, on the other hand, can be regarded as a kind of workplace evaluation.

Only the musician, singer or speaker who feels completely at ease with all fringe factors will also praise the acoustical quality. The main factors judged here are the sound volume and the mutual listening, which is also responsible for the intonation. An acoustically adequate response from the auditorium has to be realized for the performers so that this positive correspondence supports the overall artistic experience. The overall acoustic impression of his own work as it is perceived in the reception area plays a very subordinate role for the performer. What is important for him, however, are rehearsal conditions where the acoustics are as close as possible to those of the actual performance and acoustical criteria that depend as little as possible on the occupation density both in the audience area as well as in the platform area.

Generally, a performance room must not show any disturbing reflection characteristics like echo effects or flutter echoes. All seats have to guarantee a good audibility that is in good conformity with the auditory expectation. This requires a balanced sound of high clarity and an adequate spaciousness. Localization shifts or deviations between acoustical and visual directional impression must not occur. If the room is used as a concert hall, the spatial unity between the auditorium and the platform areas has to be maintained in order to avoid sound distortions.

Based on these considerations and well-founded, objective measurement technical examinations and subjective tests, partially in reverberation-free rooms within artificially generated sound fields, it is possible to define room-acoustical quality criteria that enable an optimum listening and acoustical experience in dependence on the usage function of the room. The wider the spectrum of usage is, the broader is the limit of the desirable reference value ranges of these criteria. Without extensive variable acoustical measures—also electronic ones—only a compromise brings about a somewhat satisfactory solution. It stands to reason that this compromise can only be as good as the degree in which the room-acoustical requirements coincide with it.

A precondition for an optimum room-acoustical design of auditoriums and concert halls is the very early coordination in the planning phase. The basis here is the establishment of the room's *primary structure* according to its intended use (room shape, volume, topography of the spectators' and the platform areas). The *secondary structure* that decides the design of the structures on the walls and ceilings as well as their acoustic effectiveness has to be worked out on this basis. A planning methodology for guaranteeing the room-acoustical functional and quality assurance of first-class concert halls and auditoriums as well as rooms with a complicated primary structure is reflected in the application of simulation tests by means of mathematical and physical models (see also Chapter 35 and Section7.3.1).

7.2 Room-Acoustical Criteria, Requirements

The acoustical evaluation by listeners and actors of the acoustical playback-quality of a signal that is emitted from a natural acoustic source or via electroacoustical devices, is mostly very imprecise. This evaluation is influenced by existing objective causes like disturbing climatic, seating, and visibility conditions as well as by subjective circumstances like, for instance, the subjective attitude and receptiveness towards the content and the antecedents of the performance. Very differentiated is the subjective rating of music, where the term *good acoustics* is defined, depending on the genre, as a sufficient sound volume, a good time and register clarity of the sound, and a spaciousness that meets the composition. Timbre changes that deviate from the natural-timbre of the acoustic sources and from the usual distance dependence (high-frequency sounds are less effective at a larger distance from the place of performance than at closer range) are judged as being unnatural, if traditional music is concerned. These experiences determine also the listening expectation for a very spatial and reverberating sound in a large cathedral, whereas one expects a dry sound in the open. Thus deviations from this experience are regarded as being bothersome. A listener seated in the front section of a concert hall expects a clearer sound than one seated in a rear section. On the other hand, however, he wants to enjoy an optimally balanced acoustic pattern on all seats, as he has grown up with the media and mainly postprocessed sound productions that are independent of the room, and thus acquired auditory expectations, which do not allow an evaluation of the objectively existing room.

The evaluation of speech is generally a bit easier, since optimum audibility and clear intelligibility are desired here in an atmosphere that is not influenced by the room or electroacoustical means. Perhaps with the exception of sacral rooms, the spaciousness generally does not play such an outstanding role in this regard, whereas sound volume and intelligibility are all the more important.

Numerous room-acoustical criteria were defined in order to clarify the terms applied for the subjective and objective assessment of a spoken or musical performance. In the following we have listed a relevant selection of them, in which context one should note that there is a close correlation between the individual criteria. One single optimally determined parameter may not at all be acoustically satisfactory, because another parameter influences the judgement in a negative way. For example, the optimum value range of center time and definition can only be evaluated with a subjectively correct estimated reverberation time. The guide values of the reverberance measure are valid only if the clarity measure C_{80} is in the optimal range.

On principle, the room-acoustical quality criteria can be subdivided into time and energy criteria. The main type of use—speech or music—then determines the recommendations for the guide values to be targeted. With multi-purpose halls (without available variable measures for changing the acoustics), a compromise is required that should orient itself on the main type of use.

7.2.1 Time Criteria

7.2.1.1 Reverberation Time RT_{60}

The reverberation time RT_{60} is not only the oldest, but also the most best-known room-acoustical quantity. It is the time that passes after an acoustic source in a room has been turned off until the mean steady-state sound-energy density $w(t)$ has decreased to 1/1,000,000 of the initial value w_0 or until the sound pressure has decayed to 1/1.000—i. e., by 60 dB

$$w(RT) = 10^{-6} w_0 . \tag{7-1}$$

Thus the time response of the sound energy density in reverberation[1] results as

$$\begin{aligned} w(t) &= w_0^{-6\log\frac{t}{RT}} \\ &= w_0^{-13.82\frac{t}{RT}} \end{aligned} \tag{7-2}$$

The steady-state condition is reached only after the starting time t_{st} of the even sound distribution in a room (approximately 20 sound reflections within 10 ms) after the starting time of the excitation[2]

$$t_{st} = 1\ldots\ 2(0.17\ldots\ 0.34)\sqrt{V} \tag{7-3}$$

where,

t_{st} is in ms,

V is in m³ (ft³).

The defined drop of the sound pressure level of 60 dB corresponds roughly to the dynamic range of a large orchestra.[3] The listener, however, can follow the decay process only until the noise level in the room becomes perceptible. This subjectively assessed parameter *reverberation time duration* thus depends on the excitation level as well as on the noise level.

The required evaluation dynamic range is difficult to achieve even with objective measuring, especially in the low frequency range. Therefore, the reverberation time is determined by measuring the sound level decay in a range from −5 dB to −35 dB and then defined as $RT_{30\text{ dB}}$ (also RT_{30}). The initial reverberation time (IRT according to Atal,[2] $RT_{15\text{ dB}}$ between −5 dB and −20 dB) and the early decay time (EDT according to Jordan,[2] $RT_{10\text{ dB}}$ between 0 dB and −10 dB) are mostly more in conformity with the subjective assessment of the duration of reverberation, especially at low-level volumes. This also explains the fact that the reverberation time subjectively perceived in the room may vary, while the values measured objectively according to the classical definition with a dynamic range of 60 dB or 30 dB are, except admissible fluctuations, generally independent of the location.

Serving as a single indicator for the principal characterization of the room in an occupied or unoccupied state, the reverberation time is used as the mean value between the two octave bandwidths 500 Hz and

1000 Hz or the four ⅓ octave bandwidths 500 Hz, 630 Hz, 800 Hz, and 1000 Hz, and referred to as the *mean reverberation time*.

The desirable convenient value of the reverberation time RT_{60} depends on the kind of performance (speech or music) and the size of the room. For auditoriums and concert halls, the desired values of the mean reverberation time for between 500 Hz and 1000 Hz with a room occupation of between 80% and 100% are given in Fig. 7-1 and the admissible frequency tolerance ranges are shown in Figs. 7-2 and 7-3. This shows that in order to guarantee a specific warmth of sound with musical performances, an increase of the reverberation time in the low frequency range is admissible (see Section 7.2.1.2), while with spoken performances a decrease of the reverberation time is desirable in this frequency range (see Section 7.2.2.9).

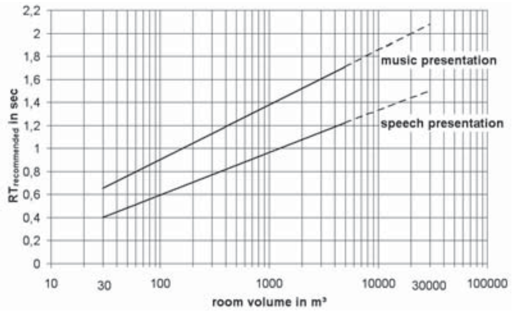


Figure 7-1. Recommended value of the mean reverberation time $RT_{recommended}$ between 500 Hz and 1000 Hz for speech and music presentations as a function of room volume *V*.

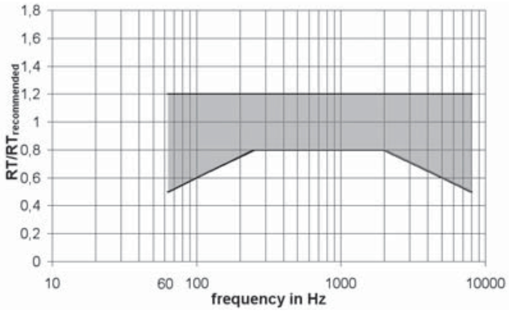


Figure 7-2. Frequency dependent tolerance range of reverberation time RT_{60} referred to $RT_{recommeded}$ for speech presentations.

The reverberation time of a room as defined by Eyring mainly depends on the size of the room and on the sound absorbing properties of the boundary surfaces and nonsurface forming furnishings:

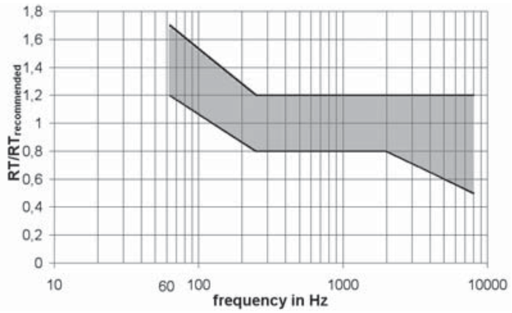


Figure 7-3. Frequency-dependent tolerance range of reverberation time RT_{60} referred to $RT_{recommended}$ for music presentations.

$$RT_{60} = 0.163^* \frac{V}{-\ln(1+\bar{a})S_{tot} + 4mV} \tag{7-4}$$

*0.049 for U.S. units

where,

RT_{60} is the reverberation time in seconds,

V is the room volume in cubic meters (cubic feet),

$\bar{a}$ is A_{tot}/S_{tot} which is the room-averaged coefficient of absorption,

A_{tot} is the total absorption surface in square meters (square feet),

S_{tot} is the total room surface in square meters (square feet),

m is the energy attenuation factor of the air in m^{-1} (see Fig. 7-4).

The correlation between the mean sound absorption coefficient and the reverberation time for different relations between room volume V and total surface S_{tot} is graphically shown in Fig. 7-5.

The total sound absorption surface of the room A_{tot} consists of the planar absorption surfaces with the corresponding partial surfaces S_n and the corresponding frequency-depending coefficient of sound absorption α_n plus the nonsurface forming absorption surfaces A_k consisting of the audience and the furnishings.

$$A_{tot} = \sum_n \alpha_n S_n + \sum_k A_k \tag{7-5}$$

For an average sound absorption coefficient of up to $\bar{a} = 0.25$, the Eq. 7-4 by Eyring[2] can be simplified by means of series expansion according to Sabine[4] to

$$RT = 0.163^* \frac{V}{A_{tot} + 4mV} \tag{7-6}$$

*0.049 for U.S. units

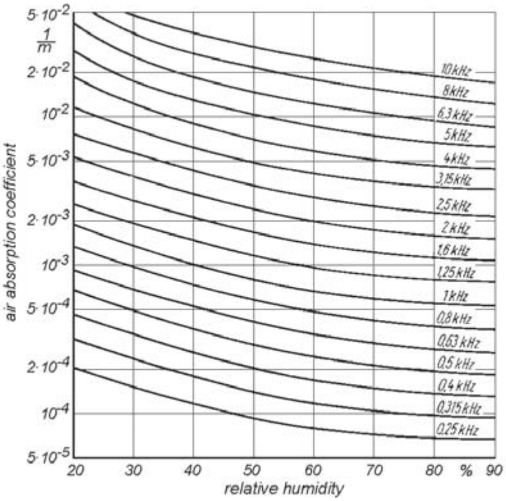


Figure 7-4. Air absorption coefficient *m* as a function of relative humidity *F*.

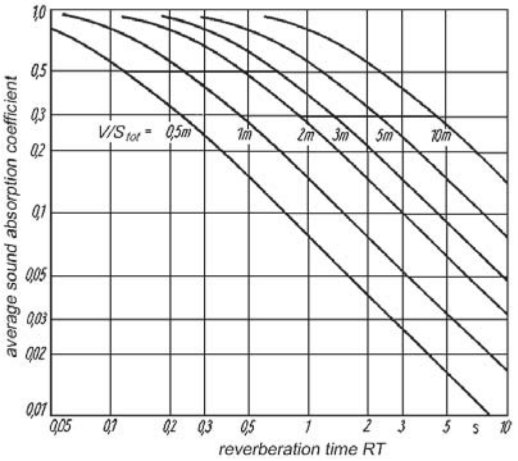


Figure 7-5. Correlation between average sound absorption coefficient and reverberation time for various ratios of room volume *V* and room surface S_{tot}.

where,

RT_{60} is the reverberation time in seconds,

V is the room volume in cubic meters (cubic feet),

A_{tot} is the total absorption surface in square meters (square feet),

m is the energy attenuation factor of the air in m^{-1}, Fig. 7-4

The correlation between the reverberation time RT_{60}, the room volume V, the equivalent sound absorption surface A_{tot}, and the unavoidable air damping m is graphically shown in Fig. 7-6.

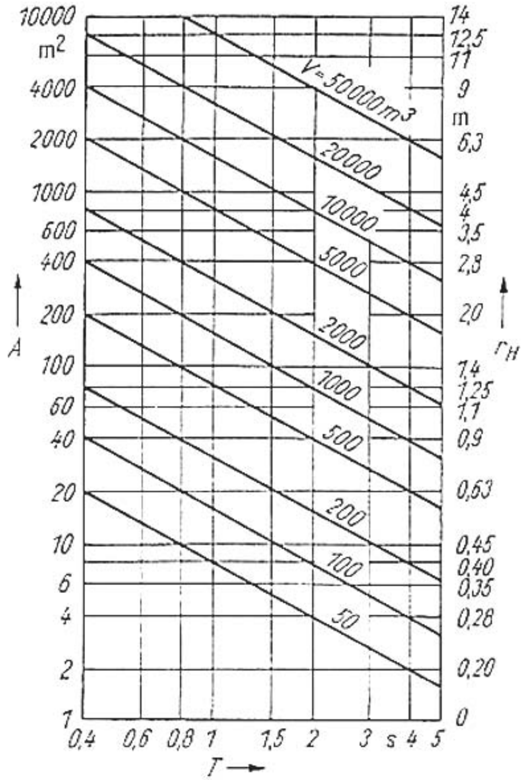


Figure 7-6. Correlation between reverberation time RT_{60}, room volume *V*, and equivalent sound absorption area *A*, according to Eq. 7-6.

The above stated frequency-dependent sound absorption coefficient has to be determined by measuring or calculation of the diffuse all-round sound incidence. Measurement is generally done in the reverberation room by using Eq. 7-6. If the sound absorption coefficient is measured by using an impedance tube (or Kundt's tube) with vertical sound incidence, the results can only be converted to the diffuse sound incidence by means of the diagrams of Morse and Bolt.[2] One can assume that the complex input impedance of the absorber is independent of the angle—i. e., if the lateral sound propagation is inhibited in the absorber (e.g. porous material with a high-specific flow resistance).

Properly speaking, the above-mentioned derivatives of the reverberation time from the sound absorption in the room are only valid for approximately cube-shaped rooms with an even distribution of the sound absorbing surfaces within the room. With room shapes deviating heavily from a square or a rectangle, or in case of a necessary one-sided layout of the absorbing audience area, these factors also have a decisive effect on the reverberation time. With the same room volume and the same equivalent sound absorption surface in the room, inclining the side wall surfaces towards the room's ceiling or towards the sound absorbing audience area results in deviations of the measured reverberation time of up to 100%. For numerous room shapes there exist calculating methods with different degrees of exactness, for example, for cylinder-shaped rooms.[5] The cause of these differences lies mainly with the geometrical conditions of the room and their influence on the resulting path length of the sound rays determining the reverberation.

The absorbed sound power P_{ab} of a room can be derived from the ratio energy density w = sound energy W/volume V under consideration of the differential coefficient $P_{ab} = dW/dt$ representing the rate of energy decay in the room and taken from Eqs. 7-5 and 7-6.

$$P_{ab} = \frac{1}{4}cwA \qquad (7\text{-}7)$$

where,
c is the sound velocity.

In steady-state, the absorbed sound power is equal to the power P fed into the room. This results in the average sound energy density w_r in the diffuse sound field of the room as

$$w_r = \frac{4P}{cA} \qquad (7\text{-}8)$$

While the sound energy density w_r in the diffuse sound field is approximately constant, the direct sound energy and thus also its density w_d decreases at close range to the source with the square of the distance r from the source, according to

$$w_d = \frac{P}{c} \times \frac{1}{4\pi r^2} \qquad (7\text{-}9)$$

Strictly speaking, this is valid only for spherical acoustic sources;[6] given a sufficient distance it can be applied, however, to most practically effective acoustic sources.

For the sound pressure in this range of predominantly direct sound, this results in a decline with $1/r$. (Strictly speaking, this decline sets in only outside of an interference zone, the near field. The range of this near field is of the order of the dimensions of the source and 0.4 m away from its center.)

If the direct sound and the diffuse sound energy densities are equal ($w_d = w_r$), Eqs. 7-8 and 7-9 can be equated, which means it is possible to determine a specific distance from the source or the reverberation radius (critical distance for omnidirectional sources) r_H. With a spherical acoustic source there is

$$\begin{aligned} r_H &= (0.3^*)\sqrt{\frac{A}{16\pi}} \\ &\approx (0.3^*)\sqrt{\frac{A}{50}} \\ &\approx 0.041(0.043^*)\sqrt{A} \\ &\approx 0.057(0.01^*)\sqrt{\frac{V}{RT}} \end{aligned} \qquad (7\text{-}10)$$

* for U. S. units

where,
r_H is in meters or feet,
A is in square meters or square feet,
V is in cubic meters of cubic feet,
RT_{60} is in seconds.

With a directional acoustic source (loudspeaker, sound transducer), this distance is replaced by the critical distance r_R

$$r_R = \Gamma(\vartheta)\sqrt{\gamma(r_H)} \qquad (7\text{-}11)$$

where,
$\Gamma(\vartheta)$ is the angular directivity ratio of the acoustic source—the ratio between the sound pressure that is radiated at the angle θ against the reference axis and the sound pressure that is generated on the reference axis at the same distance, in other words, the polars,
γ is the front-to-random factor of the acoustic source.

7.2.1.2 Bass Ratio (BR) (Beranek)

Besides the reverberation time RT_{60} at medium frequencies, the frequency response of the reverberation time is of great importance, especially at low frequencies, as compared to the medium ones. The bass ratio—i. e., the ratio between the reverberation times at octave center frequencies of 125 Hz and 250 Hz and octave center frequencies of 500 Hz and 1000 Hz (average reverberation time)—is calculated basing on the following relation:[7]

$$BR = \frac{RT_{125Hz} + RT_{250Hz}}{RT_{500Hz} + RT_{1000Hz}} \qquad (7\text{-}12)$$

For music, the desirable bass ratio is $BR \approx 1.0–1.3$. For speech, on the other hand, the bass ratio should at most have a value of $BR \approx 0.9–1.0$.

7.2.2 Energy Criteria

According to the laws of system theory, a room can be acoustically regarded as a linear transmission system that can be fully described through its impulse response $h(t)$ in the time domain. If the unit impulse $\delta(t)$ is used as an input signal, the impulse response is linked with the transmission function in the frequency domain through the Fourier transform

$$\underline{G}(\omega) = F\{h(t)\} \tag{7-13}$$

where,

$$h(t) = F^{-1}\{\underline{G}(\omega)\}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \underline{G}(\omega)^{(j\omega t)\mathrm{d}\omega}.$$

As regards the measuring technique, the room to be examined is excited with a very short impulse (delta unit impulse) and the impulse response $h(t)$ is determined at defined locations in the room, Fig. 7-7.

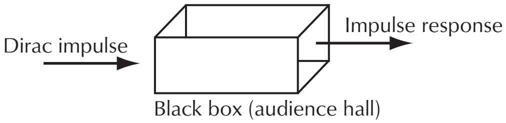


Figure 7-7. Basic solution of signal theory for identification of an unknown room.

Here, the impulse response contains the same information as a quasi-statically measured transmission frequency response.

Generally, the time responses of the following sound-field-proportionate factors (so-called reflectograms) are derived from measured or calculated room impulse responses $h(t)$

Sound pressure: $p(t) \approx h(t)$ (7-14)

Sound energy density: $w(t) \approx h^2(t)$ (7-15)

Ear-inertia weighted sound intensity:

$$J_{\tau_0} \approx \int_0^{\tau} h^2(t')^{\left(\frac{T'-t}{\tau_0}\right)\mathrm{d}t'} \tag{7-16}$$

where,
τ_0 is 35 ms.

Sound energy: $W(t) \approx \int_0^{\tau} h^2(t')\mathrm{d}t'$ (7-17)

Basic reflectogram figures are graphically shown in Fig. 7-8.

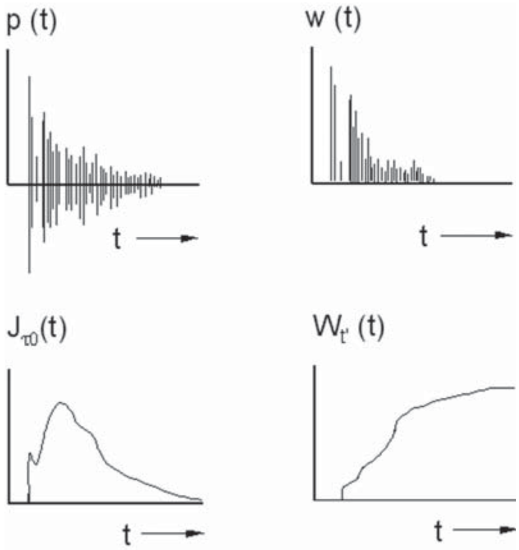


Figure 7-8. Behavior of sound field quantity versus time (reflectograms) for sound pressure $p(t)$, sound energy density $w(t)$, ear-inertia weighted sound intensity $J_{\tau 0}(t)$ and sound energy $W_{t'}(t)$.

In order to simplify the mathematical and measuring-technical correlations, a sound-energy-proportional factor is defined as *sound energy component E_f*. Being a proportionality factor of the sound energy, this factor shows the dimension of an acoustical impedance and is calculated from the sound pressure response $p(t)$.

$$\textit{Sound energy component } E_{t'} = \int_0^{t'} p^2(t)\mathrm{d}t \tag{7-18}$$

where,
t' is in ms.

For determining a sound-volume-equivalent energy component, t' has to be set to equal ∞. In practical rooms of medium size, $t' \approx 800$ ms is sufficient.

For measuring of all speech-relevant room-acoustical criteria, an acoustic source with the frequency-dependent directivity pattern of a human speaker has, on principle, to be used for exciting the sound field, while with musical performances it is sufficient to use a nondirectional acoustic source for the first approximation.

The majority of the room-acoustical quality criteria is based on the monaural, directionally unweighted assessment of the impulse response. Head-related binaural criteria are still an exception. The influence of the sound-incidence direction of early initial reflections on room-acoustical quality criteria is principally known. Since subjective evaluation criteria are still missing to a large extent, this may also be generally disregarded when measuring or calculating room impulse responses. For the determination of most of the relevant criteria, the energetic addition of the two ear signals of an artificial head is sufficient.

Just like the directional dependence, the frequency dependence of the room-acoustical energy criteria has also not been researched in depth, so that it is generally sufficient at the moment to evaluate the octave with the center frequency of 1000 Hz.

7.2.2.1 Strength Measure (G) (P. Lehmann)

The strength measure G is the ten-fold logarithmic ratio between the sound energy components at the measuring location and those measured at 10 m distance from the same acoustic source in the free field. It characterizes the volume level

$$G = 10\log\left(\frac{E_{\infty}}{E_{\infty,10\text{ m}}}\right)\text{ dB} \tag{7-19}$$

Here, $E_{\infty,10\text{ m}}$ is the reference sound energy component existing at 10 m (32.8 ft) distance with the free sound transmission of the acoustic source.

The optimum values for musical and speech performance rooms are located between $+1\text{ dB} \le G \le +10\text{ dB}$ which means that the loudness at any given listener's seat in real rooms should be roughly equal to or twice as high as in the open at 10 m (32.8 ft) distance from the sound source.[8,9]

7.2.2.2 Sound Pressure Distribution (ΔL)

The decrease of sound pressure level ΔL in dB describes the distribution of the volume at different reception positions in comparison with a reference measuring position or also for a specific measuring position on the stage in comparison with others. If the sound energy component at the reference measuring position or for the reference measuring position on stage is labeled with E_0 and at the reception measuring position or for the measuring position on stage with E, one calculates a sound pressure level distribution ΔL

$$\Delta L = 10\log\left(\frac{E_{\infty}}{E_{\infty,0}}\right)\text{ dB}\,. \tag{7-20}$$

It is advantageous for a room if ΔL for speech and music is in a range of $0\text{ dB} \ge \Delta L \ge -5\text{ dB}$.

7.2.2.3 Interaural Cross-Correlation Coefficient (IACC)

The IACC is a binaural, head-related criterion and serves for describing the equality of the two ear signals between two freely selectable temporal limits t_1 and t_2. In this respect, however, the selection of these temporal limits, the frequency evaluation as well as the subjective statement, are not clarified yet. In general, one can examine the signal identity for the initial reflections ($t_1 = 0$ ms, $t_2 = 80$ ms) or for the reverberation component ($t_1 \ge t_{st}$, $t_2 \ge RT_{60}$ [see Section 7.2.1.1]). The frequency filtration should generally take place in octave bandwidths of between 125 Hz and 4000 Hz.

The standard interaural cross-correlation function IACF[10,11] is defined as

$$IACF_{t_1t_2}(\tau) = \frac{\int_{f_1}^{f_2} p_L(t)\times p_R(t+\tau)\text{d}t}{\sqrt{\left[\int_{f_1}^{f_2} p_L^{\,2}(t)\text{d}t \times \int_{f_1}^{f_2} p_R^{\,2}(t)\text{d}t\right]}} \tag{7-21}$$

where,

$p_L(t)$ is the impulse response at the entrance to the left auditory canal,

$p_R(t)$ is the impulse response at the entrance to the right auditory canal.

Then the interaural cross-correlation coefficient IACC is

$$IACC_{t_1t_2} = max\left|IACF_{t_1t_2}(\tau)\right|$$

for $-1\text{ ms} < \tau < +1\text{ ms}$.

7.2.2.4 Center Time (t_s) (Kürer)

For music and speech performances, the center time t_s is a reference value for spatial impression and clarity and results at a measuring position from the ratio between the summed-up products of the energy components of the arriving sound reflections and the corresponding delay times and the total energy component. It corresponds to the instant of the first moment in the squared impulse response and is thus determined according to the following ratio:

$$t_s = \frac{\sum_i t_i E_i}{E_{ges}} \qquad (7\text{-}22)$$

The higher the center time t_s is, the more spatial the acoustic impression is at the listener's position. The maximum achievable center time t_s is based on the optimum reverberation time. According to Hoffmeier,[12] there is a good correlation between center time and intelligibility of speech with a frequency evaluation of four octaves between 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz.

For music, the desirable center time t_s is $t_s \approx 70$ to 150 ms with a 1000 Hz octave, and for speech $t_s \approx 60$ to 80 ms with four octaves between 500 and 4000 Hz.

7.2.2.5 Echo Criterion (EK) (Dietsch)

If we look at the build-up function of the center time $t_s(\tau)$:

$$t_s(\tau) = \frac{\int_0^\tau |p(t)|^n (t) \mathrm{d}t}{\int_0^\tau |p(t)|^n \mathrm{d}t} \qquad (7\text{-}23)$$

where,

the incoming sound reflection $n = 0.67$ with speech and $n = 1$ with music.

Comparing it with the difference quotient

$$EK(\tau) = \frac{\Delta t_s(\tau)}{\Delta t_E} \qquad (7\text{-}24)$$

we can discern echo distortions for music or speech when applying values of $\Delta t_E = 14$ ms for music and $\Delta t_E = 9$ ms for speech, ascertained by subjective tests.[13] The echo criterion depends on the motif. With fast and accentuated speech or music, the limit values are lower.

For 50% ($EK_{50\%}$) and 10% ($EK_{10\%}$), respectively, of the listeners perceiving this echo, the limit values of the echo criterion amount to:

- Echo perceptible with music for $EK_{50\%} \geq 1.8$; $EK_{10\%} > 1.5$ for two octave bands 1 kHz and 2 kHz mid frequencies.
- Echo perceptible with speech for $EK_{50\%} \geq 1.0$; $EK_{10\%} > 0.9$ for one octave band 1 kHz.

7.2.2.6 Definition Measure C_{50} for Speech (Ahnert)

The *definition measure* C_{50} describes the intelligibility of speech and also of singing. It is generally calculated in a bandwidth of four octaves between 500 Hz and 4000 Hz from the tenfold logarithm of the ratio between the sound energy arriving at a reception measuring position up to a delay time of 50 ms after the arrival of the direct sound and the following energy:

$$C_{50} = 10\log\left(\frac{E_{50}}{E_\infty - E_{50}}\right) \text{ dB} \qquad (7\text{-}25)$$

A good intelligibility of speech is generally given when $C_{50} \geq 0$ dB.

The frequency-dependent definition measure C_{50} should increase by approximately 5 dB with octave center frequencies over 1000 Hz (starting with the octave center frequencies 2000 Hz, 4000 Hz, and 8000 Hz), and decrease by this value with octave center frequencies below 1000 Hz (octave center frequencies 500 Hz, 250 Hz, and 125 Hz).

According to *Höhne* and *Schroth,*[14] the limits of the perception of the difference of the definition measure are at $\Delta C_{50} \approx \pm 2.5$ dB.

An equivalent, albeit less used criterion, is the degree of definition D, also called D_{50}, that results from the ratio between the sound energy arriving at the reception measuring position up to a delay time of 50 ms after the arrival of the direct sound and the entire energy (given in %) is

$$D = \frac{E_{50}}{E_\infty} \qquad (7\text{-}26)$$

The correlation with the definition measure C_{50} is determined by the equation

$$C_{50} = 10\log\left(\frac{D_{50}}{1-D_{50}}\right)\text{ dB} \qquad (7\text{-}27)$$

One should thus strive for an intelligibility of syllables of at least 85%, $D = D_{50} \geq 0.5$, or 50%.

7.2.2.7 Speech Transmission Index (STI) (Houtgast, Steeneken)

The determination of the STI values is based on measuring the reduction of the signal modulation between the location of the sound source—e.g., on stage—and the reception measuring position with octave center frequencies of 125 Hz up to 8000 Hz. Here Steeneken and Houtgast[15] have proposed to excite the room or open space to be measured with a special modulated noise and then to determine the reduced modulation depth.

The authors proceeded on the assumption that not only reverberation and noise reduce the intelligibility of speech, but generally all external signals or signal changes that occur on the path from source to listener. For ascertaining this influence they employ the *modulation transmission function* (MTF) for acoustical purposes. The available useful signal S (signal) is put into relation with the prevailing interfering signal N (noise). The determined *modulation reduction factor* $m(F)$ is a factor that characterizes the interference with speech intelligibility

$$m(F) = \frac{1}{\sqrt{1+\left(\frac{2\pi F(RT)}{13.8}\right)^2}} \times \frac{1}{1+10^{-\left(\frac{SNR}{10dB}\right)}} \qquad (7\text{-}28)$$

where,

F is the modulation frequency in hertz,

RT_{60} is the reverberation time in seconds,

SNR is the signal/noise ratio in dB.

To this effect one uses modulation frequencies from 0.63 Hz to 12.5 Hz in third octaves. In addition, the modulation transmission function is subjected to a frequency weighting (WMTF—weighted modulation transmission function), in order to achieve a complete correlation to speech intelligibility. In doing so, the modulation transmission function is divided into 7 octave bands, which are each modulated with the modulation frequency.[14] This results in a matrix of 7 × 14 = 98 modulation reduction factors, m_i.

The (apparent) effective signal-noise ratio X can be calculated from the modulation reduction factors m_i

$$X_i = 10\log\left(\frac{m_i}{1-m_i}\right)\text{ dB} \qquad (7\text{-}29)$$

These values will be averaged and for the seven octave bands the Modulation Transfer Indices $MTI = (X_{average} + 15)/30$, are calculated. After a frequency weighting in the seven bands (partially separated for male or female speech) you obtain the Speech Transmission Index, *STI*.

The excitation of the sound field is done by means of a sound source having the directivity behavior of a human speaker's mouth.

In order to render twenty years ago this relatively time consuming procedure in real-time operation, the RASTI-procedure (rapid speech transmission index) was developed from it in cooperation with the company Brüel & Kjaer.[16] The modulation transmission function is calculated for only 2 octave bands (500 Hz and 2 kHz) which are especially important for the intelligibility of speech and for select modulation frequencies—i.e., in all for only nine modulation reduction factors *mi*. However, this measure is used increasingly less.

Note: Schroeder[42] could show that the 98 modulation reduction factors $m(F)$ may also be derived from a measured impulse response

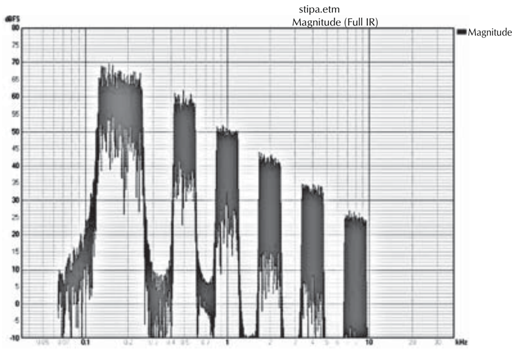
$$m(F) = \frac{\int_0^{\infty} h^2(t)e^{-j2\pi Ft}\,\mathrm{d}t}{\int_0^{\infty} h^2(t)\,\mathrm{d}t} \qquad (7\text{-}30)$$

This is done now with modern computer-based measurement routines like MLSSA, EASERA, or Win-MLS.

A new method to estimate the speech intelligibility measures an impulse response and derives STI values with the excitation with a modulated noise. The frequency spectrum of this excitation noise is shown in Fig. 7-9.

You recognize ½octave band noise, radiated through the sound system into the room. By means of a mobile receiver at any receiver location the STIPa values can be determined.[8,9] Any layman may use this method and no special knowledge is needed. It is used more and more to verify the quality of emergency call systems (EN 60849),[33] especially in airports, stations or large malls.

According to the definition the STI-value is calculated by using the results of Eq. 7-29



(c)EASERA

Figure 7-9. STIPa signal in frequency presentation.

$$STI = \frac{X+15}{30} \qquad (7\text{-}31)$$

Based on the comparison of subjective examination results with a maximum possible intelligibility of syllables of 96%, the STI values are graded in subjective values for syllable intelligibility according to Table 7-1 (EN ISO 9921: Feb. 2004).

Table 7-1. Subjective Weighting for STI

Subjective Intelligibility	STI Value
unsatisfactory	0.00–0.30
poor	0.30–0.45
satisfactory	0.45–0.60
good	0.60–0.75
excellent	0.75–1.00

7.2.2.8 Articulation Loss, Alcons, with Speech (Peutz, Klein)

Peutz[17] and Klein[18] have ascertained that the articulation loss of spoken consonants *Alcons* is decisive for the evaluation of speech intelligibility in rooms. Starting from this discovery they developed a criterion for the determination of intelligibility:

$$Alcons \approx 0.652\left(\frac{r_{LH}}{r_H}\right)^2 RT_{60}\ \% \qquad (7\text{-}32)$$

where,

r_{LH} is the distance sound source-listener,

r_H is the reverberation radius or, in case of directional sound sources, critical distance r_R,

RT_{60} is the reverberation time in seconds.

From the measured room impulse response one can determine *Alcons* according to Peutz,[17] if for the direct sound energy one applies the energy after about 25 ms to 40 ms (default 35 ms), and for the reverberation energy the residual energy after 35 ms

$$Alcons \approx 0.652\left(\frac{E_\infty - E_{35}}{E_{35}}\right) RT_{60}\ \% \qquad (7\text{-}33)$$

Assigning the results to speech intelligibility yields Table 7-2.

Table 7-2. Subjective Weighting for *Alcons*

Subjective Intelligibility	*Alcons*
Ideal intelligibility	≤3%
Good intelligibility	3–8%
Satisfactory intelligibility	8–11%
Poor intelligibility	>11%
Worthless intelligibility	>20% (limit value 15%)

Long reverberation times entail an increased articulation loss. With the corresponding duration, this reverberation acts like noise on the following signals and thus reduces the intelligibility.

Fig. 7-10 shows the articulation loss, *Alcons,* as a function of the *SNR* and the reverberation time RT_{60}. The top diagram allows us to ascertain the influence of the difference L_R (diffuse sound level) – L_N (noise level) and of the reverberation time RT_{60} on the *Alcons* value, which gives $ALcons_{R/N}$. Depending on how large the *SNR* ($L_D - L_{RN}$) is, this value is then corrected in the bottom diagram in order to obtain $Alcons_{D/R/N}$. The noise and the signal level have to be entered as dBA values.

The illustration shows also that with an increase of the *SNR* to more than 25 dB, it is practically no longer possible to achieve an improved intelligibility. (In praxis, this value is often even considerably lower, since with high volumes, for example above 90 dB, and due to the heavy impedance changes in the middle ear that set on here as well as through the strong bass emphasis that occurs owing to the frequency-dependent ear sensitivity.)

7.2.2.9 Subjective Intelligibility Tests

A subjective evaluation method for speech intelligibility consists in the recognizability of clearly spoken pronounced words (so-called test words) chosen on the

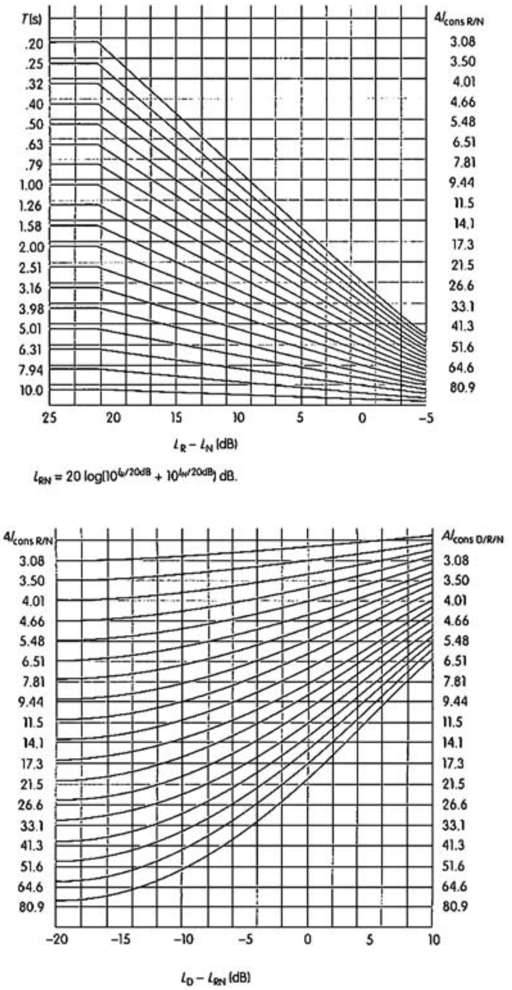


Figure 7-10. Articulation loss *Alcons* as a function of the level ratio between diffuse sound L_R and direct-sound level L_D, reverberation time RT_{60} and noise level L_N.

basis of the word-frequency dictionary and a language-relevant phoneme distribution. In German intelligibility test *logatoms* (monosyllable consonant-vowel-groups that do not readily make sense, so that a logical supplementation of logatoms that were not clearly understood during the test is not possible—e.g., *grirk*, *spres*) are used for exciting the room. In English-speaking countries, however, test words as shown in Table 7-3 are used.[19] There are between 200 and 1000 words to be used per test. The ratio between correctly understood words (or logatoms or sentences) and the total number read yields the word or syllable or sentence intelligibility V rated in percentages. The intelligibility of words V_W and the intelligibility of sentences V_S can be derived from Fig. 7-11.

Table 7-3. Examples of English Words Used in Intelligibility Tests

aisle
barb
barge
bark
baste
bead
beige
boil
choke
chore
cod
coil
coon
coop
cop
couch
could
cow
dale
dame
done
dub
feed
feet
file
five
foil
fume
fuse
get
good
guess
hews
hive
hod
hood
hop
how
huge
jack
jam
law
lawn
lisle
live
loon
loop
mess
met
neat
need
oil
ouch
paw
pawn
pews
poke
pour
pure
rack
ram
ring
rip
rub
run
sale
same
shod
shop
should
shrill
sip
skill
soil
soon
soot
soup
spill
still
tale
tame
toil
ton
trill
tub
vouch
vow
whack
wham
woe
woke
would
yaw
yawn
yes
yet
zing
zip

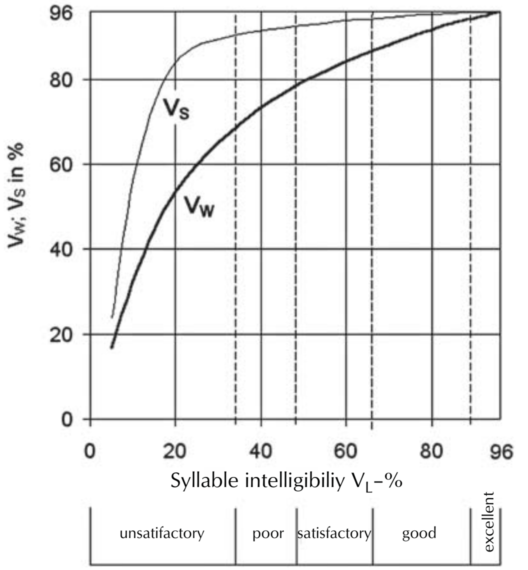


Figure 7-11. Assessment of the quality of speech intelligibility as a function of syllable intelligibility V_L, word intelligibility V_W, and sentence intelligibility V_S.

Table 7-4 shows the correlation between the intelligibility values and the ratings.

Table 7-4. Correlation between the Intelligibility Values and the Ratings

Rating	Syllable Intelligibility V_L in %	Sentence Intelligibility V_S in %	Word Intelligibility V_W in %
Excellent	90–96	96	94 –96
Good	67–90	95–96	87–94
Satisfactory	48–67	92–95	78–87
Poor	34–48	89–92	67 –78
Unsatisfactory	0–34	0–89	0–67

The results of the subjective intelligibility test are greatly influenced by speech velocity which includes the number of spoken syllables or words within the articulation time (articulation velocity) and the break time. Therefore so-called predictor sentences are often used to precede the words or logatoms that are not part of the test. These sentences consist of three to four syllables each, for example: "Mark the word...", "Please write down…," "We're going to write…." Additionally to providing a continuous flow of speech, this also serves for guaranteeing that the evaluation takes place in an approximately steady-state condition of the room.

There is a close correlation between the subjectively ascertained syllable intelligibility and room-acoustical criteria. For example, a long reverberation time reduces the syllable intelligibility[20] Fig. 7-12 owing to the occurrence of masking effects, despite an increase in loudness, see Eq. 7-8.

Quite recently, comprehensive examinations concerning the frequency dependence of speech-weighting room-acoustical criteria were conducted in order to find the influence of spatial sound coloration.[12] It was ascertained that with broadband frequency weighting between 20 Hz and 20 kHz the definition measure C_{50} (see Section 7.2.2.6) correlates very insufficiently with the syllable intelligibility. Through a frequency evaluation across three to four octaves around a center frequency of 1000 Hz, however, the influence of the sound coloration can sufficiently be taken into account. Even better results regarding the subjective weightings are provided by the frequency analysis, if the following frequency responses occur, Fig. 7-13.

As the definition declines with rising frequency due to sound coloration, the intelligibility of speech is also low (bad intelligibility → 3). This includes also the definition responses versus frequency with a maximum value at 1000 Hz, poor intelligibility → 4 in Fig. 7-13.

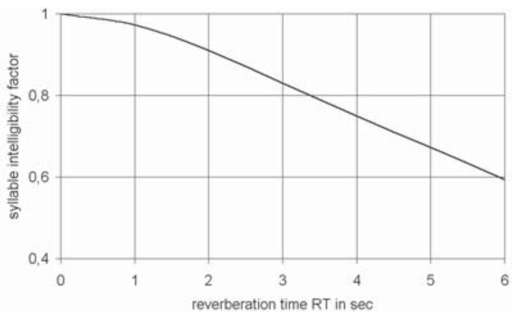


Figure 7-12. Syllable intelligibility factor as a function of reverberation time.

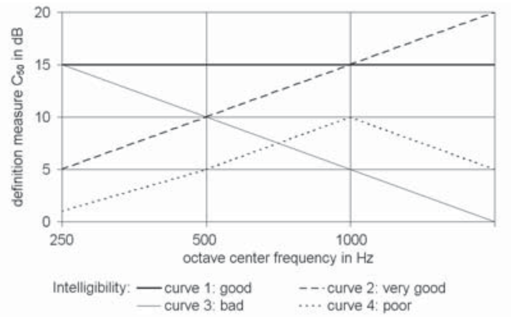


Figure 7-13. Correlation between attainable intelligibility and frequency dependence of the definition measure C_{50}.

The definition responses versus rising frequency to be aimed at for room-acoustical planning should either be constant (good intelligibility → 1) or increasing (very good intelligibility → 2). With regard to auditory psychology, this result is supported by the importance for speech intelligibility of the consonants situated in this higher-frequency range.

The determination of speech intelligibility through the definition measure C_{50} can easily lead to faulty results as the mathematical integration limit of 50 ms is not a jump function with regards to intelligibility without knowledge of the surrounding sound reflection distribution.

The best correlation with the influence of the spatial sound coloration exists between the subjective speech intelligibility and the center time t_S (see Section 7.2.2.4) with a frequency weighting between the octave of 500 Hz to the octave of 4000 Hz. According to Hoffmeier,[12] the syllable intelligibility V measured at the point of detection is then calculated as

$$V = 0.96 \times V_{sp} \times V_{SNR} \times V_R \qquad (7\text{-}34)$$

where,

V_{sp} is the influence factor of the sound source (trained speaker $V_{sp} = 1$, untrained speaker $V_{sp} \approx 0.9$),

V_{SNR} is the influence factor of the useful level (speech level) L_x and of the disturbance level L_{st} according to Fig. 7-14,[6]

$$V_R = -6 \times 10^{-6}\left(\frac{t_s}{ms}\right)^2 - 0.0012\left(\frac{t_s}{ms}\right) + 1.0488$$

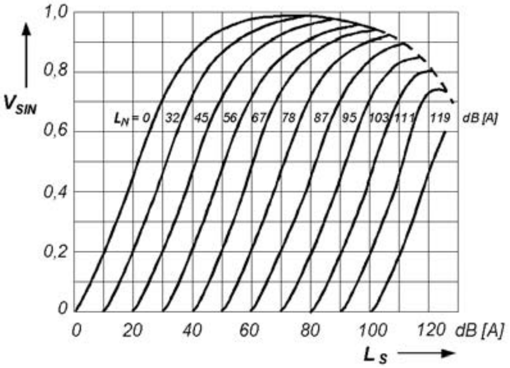


Figure 7-14. Syllable intelligibility factor $V_{S/N}$ as a function of speech sound pressure level L_S and noise pressure level L_N.

The correlation shown in Fig. 7-15 can also be derived between articulation loss and syllable intelligibility. For low reverberation times, syllable intelligibility is almost independent of the articulation loss. An inverse correlation behavior sets in only with increasing reverberation time. It is evident that with the usual *Alcons* values between 1% and 50%, syllable intelligibility can take on values between 68% and 93% (meaning a variation of 25%) and that for an articulation loss <15% (the limit value of acceptable intelligibility) the syllable intelligibility V_S reaches always, independently of the reverberation time, values over 75% which corresponds roughly to a definition measure of C_{50}> –4 dB.

This correlation can also be seen in Fig. 7-16, which shows the correlation between measured RASTI-values and articulation loss *Alcons*. One sees that acceptable articulation losses of *Alcons* <15% require RASTI values in the range from 0.4 to 1 (meaning between satisfactory and excellent intelligibility). Via the equation

$$RASTI = 0.9482 - 0.1845\ln(Alcons) \tag{7-35}$$

it is also possible to establish an analytical correlation between the two quantities. In good approximation this relationship may be used not only for RASTI but for STI as well.

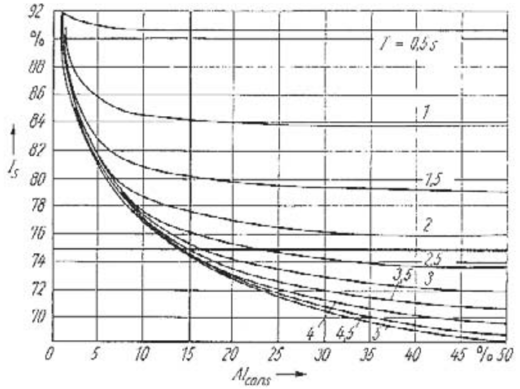


Figure 7-15. Syllable intelligibility I_S as a function of the articulation loss *Alcons*. Parameter: reverberation time RT_{60}. Preconditions: approximate statistical reverberation behavior; signal-to-noise ratio (*SNR)* ≥ 25 dB.

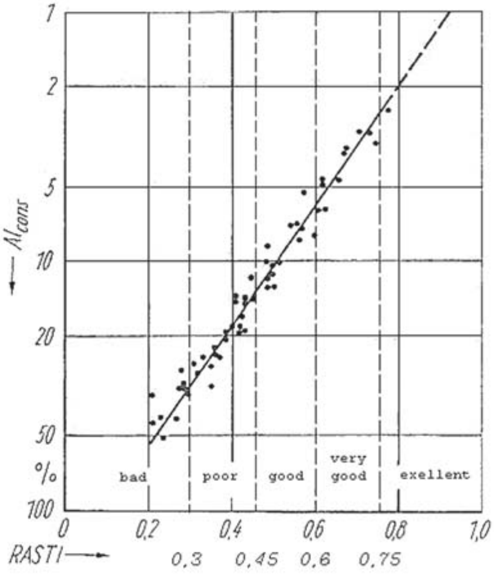


Figure 7-16. Relationship between *Alcons* values and *RASTI* values.

7.2.2.10 Clarity Measure (C80) for Music (Abdel Alim)

The clarity measure C_{80}[21] describes the temporal transparency of musical performances (defined for an octave center frequency of 1000 Hz) and is calculated from the tenfold logarithm of the ratio between the sound energy arriving at a reception measuring position up to 80 ms after the arrival of the direct sound and the following sound energy.

$$C_{80} = 10\log\left(\frac{E_{80}}{E_{\infty} - E_{80}}\right)\text{ dB} \tag{7-36}$$

The value for a good clarity measure C_{80} depends strongly on the musical genre. For romantic music, a range of approximately $-3\text{ dB} \le C80 \le +4\text{ dB}$ is regarded as being good, whereas classic and modern music will allow values up to + 6 to +8 dB.

According to Höhne and Schroth,[14] the perception limit of clarity measure differences is about $\Delta C_{80} \approx \pm 3.0$ dB.

According to Reichardt et al.,[22] there is an analytical correlation between the clarity measure C_{80} and the center time t_S, as given by

$$\begin{aligned} C_{80} &= 10.83 - 0.95t_S \\ t_S &= 114 - 10.53C_{80} \end{aligned} \tag{7-37}$$

where,

C_{80} is in dB,

t_S is in ms.

This correlation is graphically depicted in Fig. 7-17.

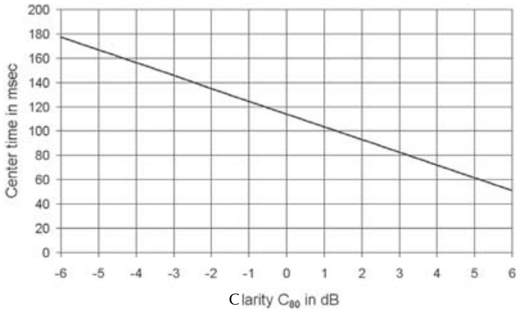


Figure 7-17. Center time t_S as a function of the clarity measure C_{80}.

7.2.2.11 Sound Coloration Measures (KT) and (KH) for Music (Schmidt)

The sound coloration measures[23] evaluate the volume-equivalent energy fractions of the room impulse response of low- and high-frequency components (K_T, octave around 100 Hz and K_H, octave around 3150 Hz, respectively) related to a medium-frequency range in an octave bandwidth of 500 Hz.

$$K_T = 10\log\left(\frac{E_{\infty,\,100Hz}}{E_{\infty,\,500Hz}}\right)\text{dB} \tag{7-38}$$

$$K_H = 10\log\left(\frac{E_{\infty,\,3150\text{Hz}}}{E_{\infty,\,500\text{Hz}}}\right)\text{dB} \tag{7-39}$$

The measures correlate with the subjective impression of the spectral sound coloration conditioned by the acoustical room characteristics. Optimum values are $K_{T;H} = -3$ to +3 dB.

7.2.2.12 Spatial Impression Measure (R) for Music (U. Lehmann)

The spatial impression measure $R^{24,25}$ consists of the two components spaciousness and reverberance. The spaciousness is based on the ability of the listener to ascertain through more or less defined localization that a part of the arriving direct sound reaches him not only as direct sound from the sound source, but also as reflected sound from the room's boundary surfaces (the perception of envelopment in music). The reverberance is generated by the nonstationary character of the music that constantly generates build-up and decaying processes in the room. As regards auditory perception, it is mainly the decaying process that becomes effective as reverberation. Both components are not consciously perceived separately, their mutual influencing of the room is very differentiated.[26] Among the energy fractions of the sound field that increase the spatial impression are the sound reflections arriving after 80 ms from all directions of the room as well as sound reflections between 25 ms and 80 ms, that are geometrically situated outside a conical window of ±40°, whose axis is formed between the location of the listener and the center of the sound source. Thus all sound reflections up to 25 ms and the ones from the front of the above-mentioned conical window have a diminishing effect on the spatial impression of the room. The tenfold logarithm of this relation is then defined as the spatial impression measure R in dB.

$$R = 10\log\left[\frac{(E_{\infty} - E_{25}) - (E_{80R} - E_{25R})}{E_{25} + (E_{80R} - E_{25R})}\right]\text{ dB} \tag{7-40}$$

where,

E_R is the sound energy fraction measured with a directional microphone (beaming angle ±40° at 500 Hz to 1000 Hz, aimed at the sound source).

One achieves a mean (favorable) room impression if the spatial impression measure R is within a range of approximately $-5\text{ dB} \le R \le +1\text{ dB}$.

Spatial impression measures below −5 dB up to −10 dB are referred to as being less spatial, others between +1 dB up to +7 dB as very spatial.

7.2.2.13 Lateral Efficiency (LE) for Music (Jordan), (LF) (Barron) and (LFC) (Kleiner)

For the subjective assessment of the apparent extension of a musical sound source—e.g., on stage—the early sound reflections arriving at a listener's seat from the side are of eminent importance as compared with all other directions. Therefore the ratio between the laterally arriving sound energy components and those arriving from all sides, each within a time of up to 80 ms, is determined and its tenfold logarithm calculated therefrom.

If one multiplies the arriving sound reflections with $\cos^2\sigma$, being the angle between the direction of the sound source and that of the arriving sound wave, one achieves the more important evaluation of the lateral reflections. With measurements this angle-dependent evaluation is achieved by employing a microphone with bi-directional characteristics.
Lateral Efficiency, *LE,* is

$$LE = \frac{E_{80Bi} - E_{25Bi}}{E_{80}} \tag{7-41}$$

where,

E_{Bi} is the sound energy component measured with a bidirectional microphone (gradient microphone).

The higher the lateral efficiency, the acoustically broader the sound source appears. It is of advantage if the lateral efficiency is within the range of 0.3 ≤ LE ≤ 0.8.

For obtaining a uniform representation of the energy measures in room acoustics, these can also be defined as lateral efficiency measure 10log *LE*. Then the favorable range is between –5 dB ≤ 10 log LE ≤ –1 dB.

According to Barron it is the sound reflections arriving from the side at a listener's position within a time window from 5 ms to 80 ms that are responsible for the acoustically perceived extension of the musical sound source (contrary to Jordan who considers a time window from 25 ms to 80 ms). This is caused by a different evaluation of the effect of the lateral reflections between 5 ms and 25 ms.

The ratio between these sound energy components is then a measure for the lateral fraction *LF*:

$$LF = \frac{E_{80Bi} - E_{5Bi}}{E_{80}} \tag{7-42}$$

where,

E_{Bi} is the sound energy component, measured with a bidirectional microphone (gradient microphone).

It is an advantage if *LF* is within the range of 0.10 ≤ *LF* ≤ 0.25, or, with the logarithmic representation of the lateral fraction measure 10log*LF*, within –10 dB ≤ 10log *LF* ≤ –6 dB.

Both lateral efficiencies *LE* and *LF* have in common that, thanks to using a gradient microphone, the resulting contribution of a single sound reflection to the lateral sound energy behaves like the square of the cosine of the reflection incidence angle, referred to the axis of the highest microphone sensibility.[27] Kleiner defines, therefore, the lateral efficiency coefficient *LFC* in better accordance with the subjective evaluation, whereby the contributions of the sound reflections vary like the cosine of the angle.

$$LFC = \frac{\int_{5}^{80} \left|p_{Bi}(t) \times p(t)\right| \mathrm{d}t}{E_{80}} \tag{7-43}$$

7.2.2.14 Reverberance Measure (H) (Beranek)

The reverberance measure describes the reverberance and the spatial impression of musical performances. It is calculated for the octave of 1000 Hz from the tenfold logarithm of the ratio between the sound energy component arriving at the reception measuring position as from 50 ms after the arrival of the direct sound and the energy component that arrives at the reception position within 50 ms.

$$H = 10\log\left(\frac{E_{\infty} - E_{50}}{E_{50}}\right)\ \mathrm{dB} \tag{7-44}$$

In contrast to the definition measure C_{50} an omnisource is used during the measurements of the reverberance measure *H*.

Under the prerequisite that the clarity measure is within the optimum range, one can define a guide value range of 0 dB ≤ *H* ≤ +4 dB for concert halls, and of –2 dB ≤ *H* ≤ +4 dB for musical theaters with optional use for concerts. A mean spatial impression is achieved if the reverberation factor *H* is within a range of –5 dB ≤ *H* ≤ +2 dB.

Schmidt[2] examined the correlation between the reverberance measure *H* and the subjectively perceived reverberation time RT_{sub}, Fig. 7-18. For a reverberance measure *H* = 0 dB, the subjectively perceived reverberation time coincides with the objectively measured reverberation time.

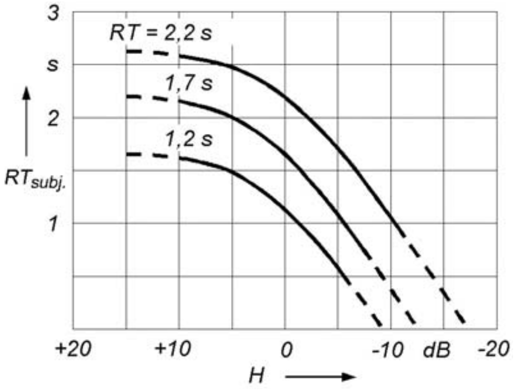


Figure 7-18. Subjectively perceived reverberation time RT_{subj} as a function of reverberance measure *H* and objective reverberation time RT_{60} as a parameter.

7.2.2.15 Register Balance Measure (BR) (Tennhardt)

With musical performances, the relation of the partial volumes of individual orchestra instrument groups between each other and to the singer is an important quality criterion for the balance (register balance) and is defined by the frequency-dependent time structure of the sound field.[28] The register balance measure *BR* between two orchestra instrument groups *x* and *y* is calculated from the A-frequency weighted volume-equivalent sound energy components of these two groups, corrected by the reference balance measure B_{xy} of optimum balance.

$$B_{Rxy} = 10\log\left(\frac{E_{\infty x}}{E_{\infty y}}\right) dB(A) + B_{xy} \tag{7-45}$$

where,
B_{xy} is in dBA.

		Group x				
		A	B	C	D	S
Group y	A	–	−5.8	1.5	0	−2.8
	B	5.8	–	7.3	5.8	3.0
	C	−1.5	−7.3	–	−1.5	−4.3
	D	0	−5.8	1.5	–	−2.8
	S	2.8	−3.0	4.3	2.8	–

Group A: String instruments,
Group B: Woodwind instruments,
Group C: Brass instruments,
Group D: Bass instruments,
Group S: Singers.

Significant differences in balance do not occur if 4 dBA < B_R < −4 dBA and if this tendency occurs binaurally.

7.3 Planning Fundamentals

7.3.1 Introduction

When planning acoustical projects one has to start out from the fundamental utilization concept envisaged for the rooms. In this respect one distinguishes between rooms intended merely for speech presentation, rooms to be used exclusively for music performances, and a wide range of multipurpose rooms.

In the following we are going to point out the most important design criteria with the most important parameters placed in front. Whenever necessary the special features of the different utilization profiles will be particularly referred to.

Strictly speaking, acoustical planning is required for all rooms as well as for open-air facilities, only the scope and the nature of the measures to be taken vary from case to case. The primordial task of the acoustician should, therefore, consist in discussing the utilization profile of the room with the building owner and the architect, but not without taking into consideration that this profile may change in the course of utilization, so that an experienced acoustician should by no means fail to pay due attention to the modern trends as well as to the utilization purposes which may arise from or have already arisen from within the environs of the new building or the facility to be refurbished, respectively.

On the one hand it is certainly not sensible for a small town to try to style the acoustical quality of a hall to that of a pure concert hall, if this type of event will perhaps take place no more than ten times a year in the hall to be built. In this case a multipurpose hall whose acoustical properties enable symphonic concerts to be performed in high quality is certainly a reasonable solution, all the more if measures of variable acoustics and of the so-called electronic architecture are included in the project.

In rooms lacking any acoustical conditioning whatsoever, on the other hand, many types of events can be performed only with certain reservations, which have to be declined from the acoustical point of view.

Table 7-5 shows the interrelation between utilization profile and effort in acoustical measures. These measures can be characterized as follows:

Table 7-5. Interrelation Between Utilization Profile and Acoustical Measures

Utilization Profile	Scope and Quality of the Acoustical Measures				
	Very High	High	Medium	Low	Very Low
Pure concert hall	x				
Pure opera house	x				
Multigenre theater		x			
Multifunctional hall, also for modern music		x			
Open air theater				x	
Club and bar areas, jazz clubs				x	
Auditoriums for speech			x		
Lecture and classrooms					x

Auditoriums, congress centers are mainly used for speech presentation. They are mostly equipped with a sound reinforcement system, but may sometimes also do without it. Music performances without sound reinforcement systems take place in a reduced style as a setting for ceremonial acts and festivities. Owing to the short reverberation time abiding their utilization concept, larger concert performances mostly require in such rooms the room-acoustical support of electroacoustical equipment (see Section 36.1).

Spoken-drama theaters serve in their classical form for speech presentation with occasional accompaniment by natural music instruments and vocalists. Apart from serving as a support for solo instruments in a music performance, utilization of electroacoustical systems is reserved almost exclusively for playing-in effects or for mutual hearing.

Multigenre theaters are gaining in the theater scene an ever-growing importance against the pure music or spoken-drama theater. The presentation of speech or music from natural sources must be possible here without compromise. While the classical music or spoken-drama theater got along with an average reverberation time of about 1 s, the trend in the planning of modern multigenre theaters tends to a somewhat longer reverberation time of up to 1.7 s with a strong portion of definition-enhancing initial sound energy and a reduced reverberance measure (less energy at the listener seat after 50 ms than within the first 50 ms). Here it may also be appropriate to make use of variable acoustics for reverberation time reduction, if, for example, electroacoustical performances (shows, pop concerts, etc.) are presented. This reverberation time reduction should be obtained by shortening the travel paths of the sound reflections rather than by sound absorption measures which tend to reduce loudness. The separation of room volumes (e.g., seats on the upper circle, reverberation chambers) leads mostly to undesirable timbre changes, unless these volumes are carefully dimensioned.

Electroacoustical systems have in multigenre theaters mostly mutual hearing and playing-in functions. Concert presentations on the stage with natural sound sources require the additional installation of a concert enclosure.

Opera houses having a large classical theater hall must be capable of transmitting speech and music presentations of natural sources in excellent acoustic quality without taking recourse to sound reinforcement. Speech is mainly delivered by singing. In room-acoustical planning of modern opera houses the parameters are therefore chosen so as to be more in line with musical requirements (longer average reverberation time of up to 1.8 s, greater spaciousness, spatial and acoustical integration of the orchestra pit in the auditorium). Electroacoustical means are used for reproducing all kinds of effects signals and for playing-in functions (e.g., remote choir or remote orchestra). This implies that the sound reinforcement system is becoming more and more an artistic instrument of the production director.

Concert presentations on the stage with natural sound sources also require the additional installation of a concert enclosure which has to form a unity with the auditorium as regards proper sound mixing and irradiation.

Multipurpose halls cover the widest utilization scope ranging from sports events to concerts. This implies that variable natural acoustics are not efficient as a planning concept, since the expenditure for structural elements generally exceeds the achievable benefit. Parting from a room-acoustical compromise solution tuned to the main intended use, with a somewhat shorter reverberation time and consequently high definition and clarity, appropriately built-in structural elements (enclosures) have to provide for the proper sound mixing required for concerts with natural music instruments, while prolongation of reverberation time as well as enhancement of spatial impression and loudness can be achieved by means of electroacoustical systems of electronic architecture (see Section 7.4).

To a greater extent sound systems are used here to cover speech amplification and the needs of modern rock and pop concerts.

Classical concert halls serve first of all for music events ranging from soloist presentations to the great symphony concert with or without choir. They are mostly equipped with a pipe organ and their room-acoustical parameters must satisfy the highest quality demands. Electroacoustical systems are used for vocal information and for mutual listening with special compositions, but are generally still ruled out for influencing the overall room-acoustical parameters.

General concert halls can be used for numerous music performances, among others for popular or pop concerts. Here it is the use of an electroacoustical sound reinforcement system that overrides the room-acoustical parameters of the hall. In accordance with the variety of events to be covered by these halls, they should be tuned to a frequency-independent reverberation time of the order of 1.2 s and feature high clarity.

Sports halls, gymnasiums have to provide an acoustical support for the mutual emotional experience. This concerns, first of all, the supporting acoustic correspondence of the spectators between themselves and the performers. Thus there are only few sound absorbing materials to be used in the spectators' areas and sound reflecting elements to be provided towards the playing field. The ceiling of the playing field should be more heavily damped so as to enable it to be used also for music events, in which case an electroacoustical sound reinforcement system is to be used. The same applies to open and partially or fully covered stadiums where sound absorption above the playing field is a natural feature with the open ones.

Show theaters are generally used only on negligibly few occasions with natural acoustics, an exception being "Singspiel" theaters with an orchestra pit. Predominantly, however, an electroacoustical sound reinforcement system is used for the functions of play-in and mutual hearing as well as half or full playback. The room-acoustical parameters of the theater room have, with this form of utilization to comply with, the electroacoustical requirements. The reverberation time should therefore not exceed a frequency-independent value of 1.4 s and the sound field should have a high diffusivity so that the electroacoustically generated sound pattern does not get distorted by the acoustics of the room.

Rooms with variable acoustics controlled by mechanical means show some positive result only in a certain frequency range, if corresponding geometric modifications of the room become simultaneously visible. The room-acoustical parameters have always to coincide with the listening experience that means they must also be perceived in a room-size and room-shape-related manner. Experimental rooms and effect realization (e.g., in a virtual stage setting of a show theater) are, of course, excluded from this mode of consideration. In theater rooms and multipurpose halls it is possible to vary the reverberation time by mechanical means within a range of about 0.5 s without detrimental effect on spatial impression and timbre. At any rate one should abstain from continuously variable acoustic parameters, house superintendent acoustics, since possible intermediate steps could lead to uncontrolled and undesirable acoustic settings.

Sacral rooms. Here we have to distinguish between classical church rooms and contemporary modern sacral buildings. With the classical rooms it is their size and importance that determine their room-acoustical parameters—e.g., a long reverberation time and an extreme spaciousness. Short reverberation times sound inadequate in such an environment. The resulting deficiency in definition, inconvenient—e.g., during the sermon—has to be compensated by providing additional initial reflections through architecturally configured reflectors, or nowadays, mostly through an electroacoustical sound system. With music presentations one has, in various frequency domains, to adapt the style of playing to the long decay time (cf. Baroque and Romanesque churches). Electroacoustical means can serve here only for providing loudness.

From the acoustical point of view, modern church buildings acquire to an increasing degree the character of multipurpose halls. Thanks to appropriately adapted acoustics and the use of sound reinforcement systems they are not only adequate for holding religious services, but can also be used as venues for concerts and conferences in good quality.

7.3.2 Structuring the Room Acoustic Planning Work

7.3.2.1 General Structure

The aim of room-acoustical planning consists of safeguarding the acoustical functionality under the envisaged utilization concepts of the auditorium for the performers as well as for the audience. With new buildings such details should be considered in the planning phase, whereas with already existing rooms an appropriate debugging should be an essential part of the refurbishment. Point of departure in this respect is a purposeful influencing control of the *primary structure* of the performance room. This concerns, among other things:

- The size of the room.
- The shape of the room.
- Functional-technological circumstances, for instance the platform or stage arrangement, the installation of balconies or galleries, lighting installations, and the arrangement of multimedia equipment.
- The topography regarding the arrangement of performers and listeners, like for instance the sloping of tiers or the proscenium area in front of the stage opening.

Based on these premises, the *secondary structure* of the room will be acoustically determined. This structure concerns essentially:

- The arrangement and distribution of frequency-dependent sound-absorbing as well as sound-reflecting faces.
- The subdivision of the surface structure for directional and diffuse sound reflections.
- The frequency-dependent effect of uneven surfaces.
- The architectural-stylistic conformation of all boundary surfaces of the room.

7.3.2.2 Room Form and Sound Form

There exists a correspondence between the shape of a room (room form) in its primary structure and the resulting sound. The term *sound form* refers in this context to the reverberation timbre which is herewith divided into its low-frequency portion (warmth) and its high-frequency portion (brilliance).

The method used for assessing the acoustical quality of concert halls is based on a paper by Beranek[46] in which one finds a list of seventy concert halls arranged in six subjective categories according to their acoustical quality. Of all these there are three halls listed in the category A+ as outstanding or superior and six halls in the category A as excellent. Eight of these are shoebox-shaped, a fact that gives rise to the question as to whether a good roomacoustical quality is linked to a rectangular shape of the room.

The subjective assessment parameters used are, on the one hand, the warmth of the sound pattern and, on the other hand, the brilliance of the same. Warmth and brilliance refer in this context mainly to the influence of the sound energy density on the lower- and the higher-frequency ranges, respectively. Questions concerning initial reflections will be left out of consideration for the time being; only the timbre in the decay process will be considered.

The criterion bass ratio, BR (Beranek[9]), provides indisputable evidence on the warmth of the sound (see Section 7.2.1.2). The desirable optimum value range for music performances is between 1.0 to 1.3. According to Beranek for rooms having a reverberation time lower than 1.8 s, it is permissible to have a bass ratio of up to 1.45.

For objective assessment of the timbre Schmidt[23] has defined the timbre measure. By analogy with the BR and the timbre measures there was an equivalent measure deduced and introduced as $TR1$ (Timbre Ratio). It is used only for evaluating comparative aspects of brilliance

$$TR1 = \frac{T_{2000\text{ Hz}} + T_{4000\text{ Hz}}}{T_{125\text{ Hz}} + T_{250\text{ Hz}}}.$$

This numerical relationship is used to evaluate timbre as the ratio of the reverberation time at high compared to low frequencies. Thus the value $TR1 > 1$ stands for a longer reverberation time at higher frequencies rather than at lower frequencies, and hence in this context higher brilliance in the sound pattern.

As regards the primary room structure of a concert hall, there are four basic forms considered: rectangle (shoebox), polygon, circle, and various trapezoidal forms.

The concert halls selected by Beranek for the categories A+ and A allow the following pairs of values to be ascertained in an occupied hall, Table. 7-6.

Table 7-7 shows that, in shoebox-shaped rooms, the brilliance is lower in comparison to rooms of polygonal primary shape. However, on the basis of the brilliance ratio $TR1$, no such significant difference can be shown between rooms of a quasi-circular ground plan (five halls) and those having diverse trapezoidal primary shapes (nine halls).

7.3.3 Primary Structure of Rooms

7.3.3.1 Volume of the Room

As a rule, the first room-acoustical criterion to be determined as soon as the intended purpose of the room has been clearly established, is the reverberation time (see Section 7.2.1.1). From Eq. 7-6 and the correlation between reverberation time, room volume and equivalent sound absorption area, graphically depicted

Table 7-6. $T_{30,mid}$, *BR,* and *TR*1 for Outstanding and Excellent Concert Halls

Room (in Alphabetical Order)	$T_{30,mid}$ in s	BR	TR1	Primary Structure
Amsterdam Concertgebouw	2.0	1.09	0.77	Rectangle
Basel, Stadt-Casino	1.8	1.18	0.74	Rectangle
Berlin, Konzerthaus	2.05	1.08	0.79	Rectangle
Boston Symphony Hall	1.85	1.03	0.78	Rectangle
Cardiff, David's Hall	1.95	0.98	0.87	Polygon
New York Carnegie Hall	1.8	1.14	0.78	Rectangle
Tokyo Hamarikyu Asahi	1.7	0.93	1.04	Rectangle
Vienna Musikvereinssaal	2.0	1.11	0.77	Rectangle
Zürich Tonhallensaal	2.05	1.32	0.58	Rectangle

Table 7-7. Brilliance Ratio for 36 Examined Concert Halls

Room shape	Number of Examined Halls	Average Value *TR*1	Confidence Range *TR*1	Limit Values *TR*1
Rectangle	12	0.75	±0.05	0.70–0.80
Polygon	10	0.91	±0.05	0.86–0.97
Circle	5	0.75	±0.16	0.59–0.91
Trapezoidal	9	0.75	±0.15	0.63–0.86

in Fig. 7-6, it becomes evident that the room volume must not fall short of a certain minimum if the desired reverberation time is to be achieved with the planned audience capacity.

For enabling a tentative estimate of the acoustically effective room size required with regard to its specific use, there serves the volume index *k*, which indicates the minimum room volume in m³/listener seat, Table 7-8. In case an auditorium is used for concert events, the volume of the concert enclosure is added to the volume of the auditorium without increasing, however, the seating capacity of the auditorium by the number of the additional performers (orchestra, choir). For theater functions the volume of the stage house behind the portal is left out of account.

The minimum required acoustically effective room volume is calculated as follows:

$$V = k \times N \tag{7-46}$$

where,

V is the acoustically effective room volume in m³,

k is the volume index according to Table 7-5 in m³/seat,

N is the seating capacity in the audience area.

If a given room is to be evaluated regarding its suitability for acoustic performances, the volume index may be useful for providing a rough estimate and simultaneously for determining the scope of additional sound-absorptive measures.

Table 7-8. Volume Index *k* versus Room Volume

No.	Main use	Volume Index *k* in m³/seat (in ft³/Seat)	Maximum Effective Room Volume with Natural Acoustics in m³ (ft³)
1	Speech performances—e.g., spoken drama, congress hall and auditorium, lecture room, room for audiovisual performances	3–6 (110–210)	5000 (180,000)
2	Music and speech performances—e.g., musical theater, multi-purpose hall, town hall	5–8 (180–280)	15,000 (550,000)
3	Music performances—e.g., concert hall	7–12 (250–420)	25,000 (900,000)
4	Rooms for oratorios and organ music	10–14 (350–500)	30,000 (1,100,000)
5	Orchestra rehearsal rooms	25–30 (900–1100)	–

If the volume index falls short of the established guide values, the desirable reverberation time cannot be achieved by natural acoustics. With very small rooms, especially orchestra rehearsal rooms, it is moreover possible that loudness results in excessive fortissimo (in a rehearsal room of 400 m³ (14,000 ft³) volume and with 25 musicians, may reach up to 120 dB in the diffuse field). In rooms of less than 100 m³ (3500 ft³) the eigenfrequency density results are insufficient.[2] This leads to a very unbalanced frequency transmission function of the room giving rise to inadmissible timbre changes.

Excessive loudness values require additional sound-absorptive measures, which may bring about too heavy a loudness reduction for low-level sound sources.

On the other hand, it is not possible to increase seating capacity and room volume just as you like, because of the increase of the equivalent sound absorption area and the unavoidable air absorption, Fig. 7-4. The attainable sound energy density in the diffuse field decreases as well as the performance loudness (see Eq. 7-8). Moreover the distances within the performance area and to the listener are dissatisfactory expanded this way. For these reasons it is possible to establish an upper volume limit for rooms without electroacoustic sound reinforcement equipment (i.e., with natural acoustics) that should not be exceeded, Table 7-2. These values depend, of course, on the maximum possible power of the sound source. By choosing Eq. 7-8 in level representation and using the reverberation time formula by Sabine,[6] one obtains the correlation between sound power level of the sound source L_W in dB and sound pressure level L_{diff} in dB in the diffuse sound field, as a function of the room parameters, volume V in m³, and reverberation time RT_{60} in s.[6]

$$L_{diff} = L_W - 10\log\frac{V}{T}\text{ dB} + 14\text{ dB*} \tag{7-47}$$

* add 29.5 dB in U.S. system.

The graphical representation of this mathematical relation is shown in Fig. 7-19. For determining the attainable sound pressure level in the diffuse sound field one can proceed from the following sound power levels L_W.[3,29,30]

Music (Mean Sound Power Level with “Forte”)	
Tail piano, open	L_W = 77–102 dB
String instruments	L_W = 77–90 dB
Woodwind instruments	L_W = 84–93 dB
Brass instruments	L_W = 94–102 dB
Chamber orchestra of 8 violins	L_W = 98 dB
Small orchestra with 31 string instruments, 8 woodwind instruments, and 4 brass instruments (without percussion)	L_W = 110 dB
Big orchestra with 58 string instruments, 16 woodwind instruments, and 11 brass instruments (without percussion)	L_W = 114 dB
Singer	L_W = 80–105 dB
Choir	L_W = 90 dB
***Speech* (Mean Sound Pressure Level with Raised to Loud Articulation)**	
Whispering	L_W = 40–45 dB
Speaking	L_W = 68–75 dB
Crying	L_W = 92–100 dB

With musical performances, for instance in forte and piano passages, perception of the dynamic range plays a decisive role for the listening experience, independently of the prevailing volume level. A sound passage or a spoken text submerging in the surrounding noise level is no longer acoustically registered and the performance considered to be faulty. The mean dynamic range of solo instruments lies with slowly played tones between 25 dB and 30 dB.[30] With orchestra music it is about 65 dB and with singers in a choir about 26 dB.

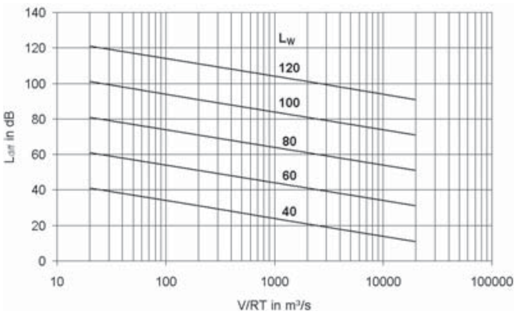


Figure 7-19. Correlation between the sound pressure level L_{diff} in the diffuse field and the sound power level L_W as a function of room volume V and reverberation time RT_{60}.

The dynamic range of a talker is about 40 dB and that of soloist singers about 50 dB. Without taking into account the timbre of a sound source, the *SNR* should generally be at least 10 dB with pianissimo or whispering.

With large room volumes and high frequencies, the increase of energy attenuation loss caused by the medium air should not be neglected. This shall be illustrated by an example: in a concert hall with a room volume of 20,000 m³ (700,000 ft³), the unavoidable air attenuation at 20°C and 40% relative humidity accounts for an additional equivalent sound absorption area which at 1000 Hz corresponds to an additional 110 persons and at 10 kHz to 5000 additional persons.

7.3.3.2 Room Shape

The shape of a room allows a wide margin of variability, since from the acoustical point of view it is not possible to define an optimum. Depending on the intended purpose, the shape implies acoustical advantages and disadvantages, but even in the spherical room of a large planetarium, it is by room-acoustical means (full absorbing surfaces) possible to achieve good speech intelligibility. Acoustically unfavorable, however, are room shapes that do not ensure an unhindered direct sound supply nor any omnidirectional incidence of energy-rich initial reflections in the reception area, as is the case, for instance, with coupled adjoining rooms and low-level audience areas under balconies or galleries of low room height.

When selecting different room shapes of equal acoustically effective volume and equal seating capacity there may result very distinct characteristics as regards the overall room-acoustical impression. A more or less pronounced inclination of the lateral boundary surfaces may produce different reverberation times.[31] In combination with a sound reflecting and not much structured ceiling layout, prolongation of the reverberation time up to a factor of two results and is especially large if long-delayed sound reflection groups are produced by side wall surfaces that are inclined outwards and upwards. But if these wall surfaces are inclined towards the sound-absorbing audience area, the shorter path lengths thus achieved may considerably reduce the reverberation time as compared to the usual calculating methods with vertical boundary surfaces.

Also with similar room shapes, different room-acoustical conditions are obtained by just varying the furnishing of the room (platform, audience areas).

All acoustically usable room shapes have in common that the unhindered direct sound and energy-rich initial reflections reach the listener. Deviations from this rule occur through direct-sound shading in the orchestra pit of an opera theater. Diffraction compensates this effect partially and the listening experience is adapted to a different sound impression which is similar to the case of unhindered sound irradiation. The initial reflections must arrive at the listener's seat within a path difference to direct sound of approximately 17 m (50 ms) for speech and 27 m (80 ms) for music performances.

Decisive for an adequate spatial impression with musical performances are, first of all, the lateral sound reflections. The more the spaciousness is supported this way, the more the orchestra sound gains, according to Meyer,[30] in volume and width. The increase of sound intensity perceptible with forte-play is thus enhanced beyond the mere loudness effect so that the subjectively perceived dynamic range is expanded. By the same token, spaciousness is subjectively enhanced by an increased loudness of the sound source.

From these general premises it is, for different arrangement patterns between performers and listeners, possible to derive universally valid guidelines for fundamental room-acoustical problems of certain typical room ground-plan layouts. In this regard one can distinguish between *purely geometrical layouts* with parallel boundary lines (rectangle, square, hexagon) on all sides, with at least two mutually slanted boundary lines (trapezoid) and generally curved boundary lines (circle, semicircle, ellipse) and *irregular layouts* with asymmetric or polygonal boundary lines.

7.3.3.2.1 Ground Plan

For obtaining lateral sound reflections, a room with a rectangular ground plan is very well suited if the performance zone is arranged at an end wall and the width of the room is in the range of 20 m (66 ft), Fig. 7-20A. This is the typical example of the linear contact in a shoebox layout of a classic concert hall (Symphony Hall Boston, Musik-vereinssaal Vienna, Konzerthaus Berlin).

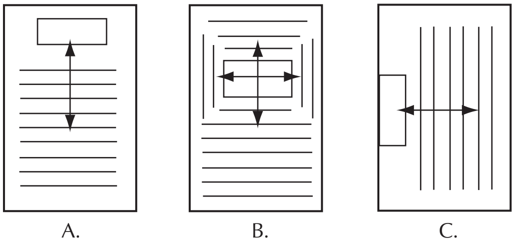


Figure 7-20. Examples of so-called arrangement patterns between performers and listeners in a room with a rectangular ground plan.

If the performance zone is shifted from the end wall towards the middle of the room, Fig. 7-20B, a circular contact may come into being as a borderline case, where audience or a choir may be arranged laterally or behind the platform. Owing to the relatively pronounced frequency-dependent directional characteristic of most sound sources (singers, high-pitched string instruments, etc.) there occur herewith, especially in the audience area arranged behind the platform, intense balance problems which may even lead to unintelligibility of the sung word and to disturbing timbre changes. On lateral seats at the side of the platform, the listening experience can be significantly impaired due to room reflections where visually disadvantaged instruments are perceived louder than instruments located at closer range. This effect is even enhanced by lateral platform boundary surfaces, whereas an additional rear sound reflection area supports sound mixing. Often these acoustical disadvantages are, however, subordinated to the more eventful visual experiences.

If the performance zone is arranged in front of a longitudinal wall, Fig. 7-20C, short-time lateral initial reflections get missed especially with broad sound sources (orchestras), whereby the mutual hearing and consequently the intonation get impaired. Soloist concerts or small orchestras (up to about six musicians) may still provide satisfactory listening conditions, if ceiling height and structure provide clarity-enhancing sound reflections. By means of a sound-reflecting rear wall combined with adjustable lateral wall elements which do not necessarily disturb the visual impression, it is possible to attain good room-acoustical conditions with not too long rooms (up to about 20 m or 66 ft). For spoken performances this way of utilization provides advantages because of the short distance to the talker, but has disadvantages due to timbre changes impairing intelligibility. As the talker chooses to speak towards the audience seats in the middle of the room, the lateral seating areas are bound to be disadvantaged on account of the frequency-dependent directional characteristic. According to Meyer,[3] the sound pressure level reduction laterally to the talker when articulating the vowels "o," "a," and "e" is 0 dB, 1 dB, and 7 dB, respectively.

A special case of the rectangular room ground plan is the square with approximately equal side lengths. Especially for multi-purpose use such halls allow the realization of diverse forms of confrontation with the audience for which good acoustical conditions are given in small rooms with about 500 seats,[29] assuming some basic principles are considered, Fig. 7-21A to C. Room variant A represents the classical linear contact ensuring a good direct-sound supply to the listeners, especially with directional sound sources (talkers, singers, directional instrumental groups). Variant B offers an acoustically good solution for sound sources of little extension (talkers, singers, chamber music groups), since a good lateral radiation into the room is given. It is true, however, that in the primary structure there is a lack of lateral sound reflections for supporting mutual hearing and intonation in the performance zone. The amphitheatrical arrangement shown in variant C is suitable for only a few kinds of performance, since apart from visual specialities there are above all acoustic balance problems to be expected. With directional sound sources—e.g., talkers and singers—the decrease of the direct sound by at least 12 dB versus the straight-ahead viewing direction produces intelligibility problems behind the sound source.

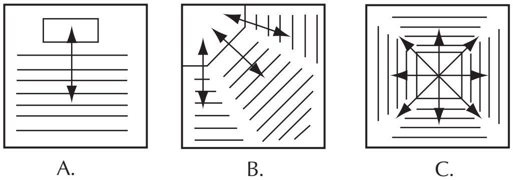


Figure 7-21. Diverse platform arrangements in a room with a square ground plan.

The trapezoidal ground plan enables, on principle, two forms of confrontation: the diverging trapezoid with the lateral wall surfaces diverging from the sound source and the converging trapezoid with the sound source located at the broad end side. The latter ground plan layout, however, is from the architectural point of view not used in its pure form, Fig. 7-22.

Variations of the first ground plan layout with a curved rear wall are designated as fan-shaped or piece of pie. The room-acoustical effect of the trapezoidal layout depends essentially on the diverging angle of the side wall surfaces. The room shape shown in Fig. 7-22A produces room-acoustical conditions, which are similarly favorable as those in a rectangular room used for music, if the diverging angle is slight. With wider diverging angles, the energy-rich initial reflections, especially from the side walls, are lacking in the whole central seating area, which is a characteristic conditioned by this primary structure. A principal comparison of the fraction of lateral sound energy produced merely by the ground plan layout is shown in Fig. 7-23. As was to be expected, the comparable, relatively narrow ground plan of the rectangular shape shows a higher lateral sound fraction than the diverging trapezoid. For spoken performances this situation is relatively uninteresting, since in most cases the lacking early lateral reflections can be compensated by early reflections from the ceiling. If the performance zone is shifted in an amphitheater-like solution to the one-third point of the room ground plan, Fig. 7-22B, this variant is suitable only for musical performances. The listeners seated behind the performance zone especially receive a very spatial, lateral-sound accentuated sound impression.

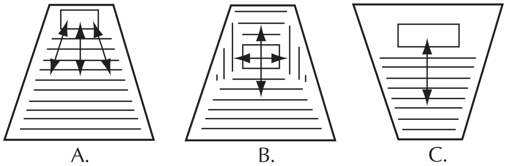


Figure 7-22. Examples of so-called arrangement patterns between performers and listeners in a room with a trapezoidal ground plan.

The room acoustically most favorable with a trapezoidal layout is that of a converging trapezoid with the performance zone located at the broad end side, Fig. 7-22C. Already without additional measures on the side of the platform, the audience areas receiving low early lateral sound energy gets reduced to a very small area in front of the sound source; almost all of the other part of the audience areas receives a strong lateral energy fraction, Fig. 7-23. Unfortunately, this room shape has only perspectives for architectural realization in combination with a diverging trapezoid as a platform area. The arrangement of so-called vineyard terraces constitutes herewith a very favorable compromise solution in which wall elements in the shape of converging trapezoids are additionally integrated in the seating area. The effective surfaces of these elements direct energy-rich initial reflections into the reception area,[32] Fig. 7-24. Examples of projects accomplished in this technique are the concert halls of the Gewandhaus in Leipzig and De Doelen in Rotterdam.[33]

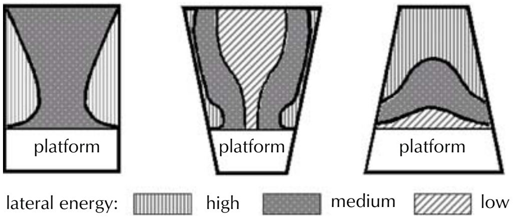


Figure 7-23. Principal portions of early lateral sound reflections in rectangular and trapezium rooms.

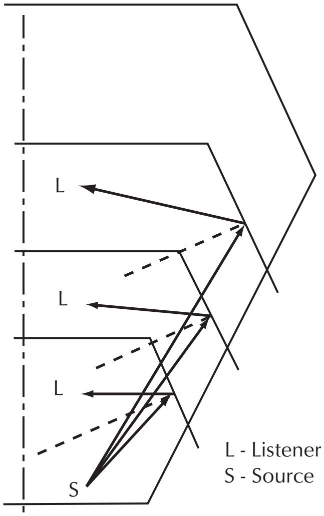


Figure 7-24. Lateral sound reflections produced by vineyard terraces.

This combination of a ground plan layout can also be realized in the shape of a hexagon which is a common application of a regular polygon ground plan. Elongated hexagons show room-acoustical properties similar to those of the combination of a diverging and a converging trapezoid or, provided the converging or diverging angle is slight, to those of a rectangular room. If the ground plan is that of a regular hexagon, Fig. 7-25, the necessary lateral sound reflections are lacking especially with musical performances. Thanks to its varied uses and the short distances between the performance and reception zones it provides, this shape is rather advantageous for congress and multipurpose halls from the acoustical point of view. The amphitheaterlike arrangement of stage and audience of Fig. 7-25D shows acoustical similarities to the rectangular variant of Fig. 7-20C. With sound sources having a pronounced directional characteristic there occur timbre and clarity problems for listeners seated at the sides and behind the platform area, which cannot be compensated by means of additional secondary structures along the walls.

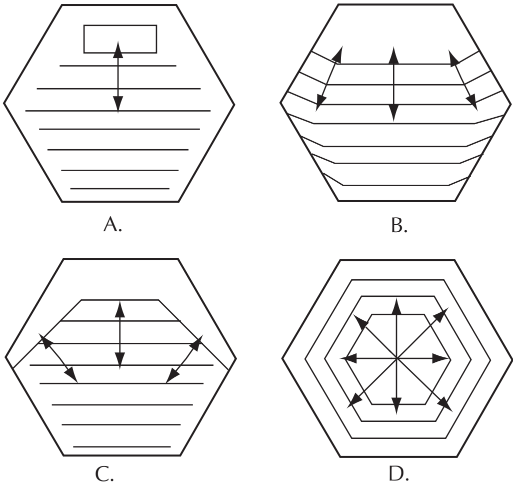


Figure 7-25. Diverse platform arrangements in a room of hexagonal shape.

Ground plans with monotonically curved boundary surfaces (circle, semicircle, Figs. 7-26A to D) produce, due to their concave configuration towards the sound source and especially if the tiers are only slightly sloped or not at all, undesirable sound concentrations. On account of the curved surfaces, the sound pressure level may, in the concentration point, even surpass that of the original sound source by 10 dB and thus become an additional disturbing sound source. The resulting wave front responses which depend on frequency, travel time, and circle diameter, are shown in Fig. 7-27.[34] One recognizes instances of migrating punctual and flat-spread sound concentration (the so-called caustic), which even after long travel times never do lead to a uniform sound distribution. Without any structuring in the vertical plane and without broadband secondary structures, rooms having a circular ground plane are acoustically suited neither for speech nor for musical performances.

With asymmetrical ground plans, Fig. 7-26E, there exists, for musical performances, the risk of a very poor correlation between the two ear signals, an effect that may give rise to an exaggerated spaciousness. Energy-rich initial reflections are to underline the visual asymmetry only as far as required for the architectural comprehension of the room, otherwise the room

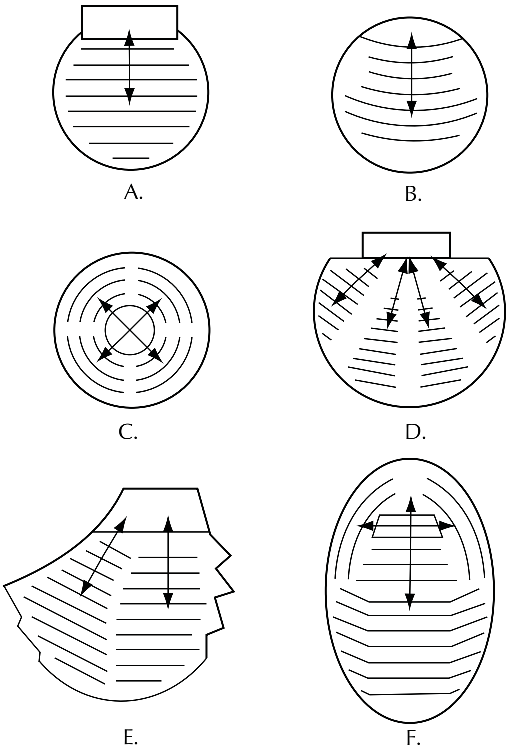


Figure 7-26. Examples of so-called arrangement patterns (confrontations) between performers and listeners in a room with curved boundary surfaces.

produces balance problems with musical performances, leaving questions regarding the arrangement of the orchestra instrument groups unsolved. Elliptical ground plans, Fig. 7-26F, are, without reflection-supporting measures, acoustically suited only for locally fixed sound sources. This general utilization is not recommended owing to the focus formation in the performance zone as well as in the audience area. This refers especially to the atrium courtyards of unstructured glass walls and plane floor in large office buildings, which are a modern architectural trend. These functionally designed entrance foyers are often used for large musical events which, however, can in no way satisfy any room-acoustical requirements.

7.3.3.2.2 Ceiling

In general, the ceiling configuration contributes little to spaciousness of the sound field, but all the more to achieving intelligibility with speech, clarity with music, volume, and guidance of reverberation-determining room reflections. For speech, the reverberation time should be dimensioned as short as possible. Therefore the ceiling should be configured in such a way that possibly each first sound reflection reaches the middle and rear audience areas, Fig. 7-28. For musical performances the mean ceiling height has to comply with the volume-index requirements. For achieving an as long as possible reverberation time, the ceiling should have its maximum height where the length or width of the room is maximal. The repeated reflection of the sound energy by the involved boundary surfaces produces long travel times, while the required slightness of energy reduction by the reflections has to be insured by a negligible sound absorption coefficient of these surfaces.[35] Thanks

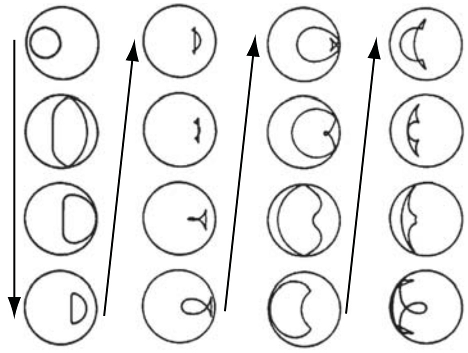


Figure 7-27. Propagation of the wave front in a circle (caustic).

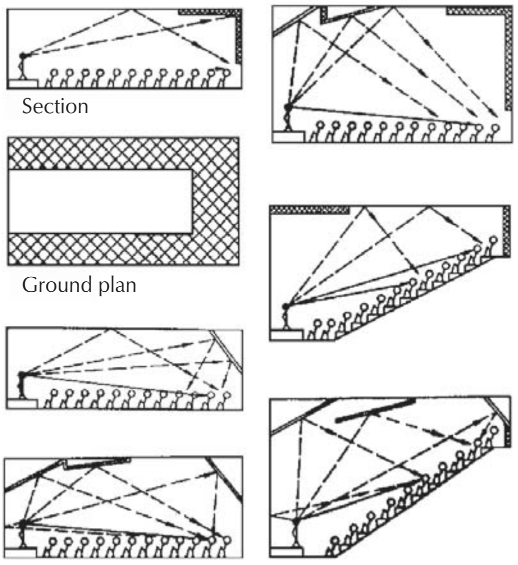


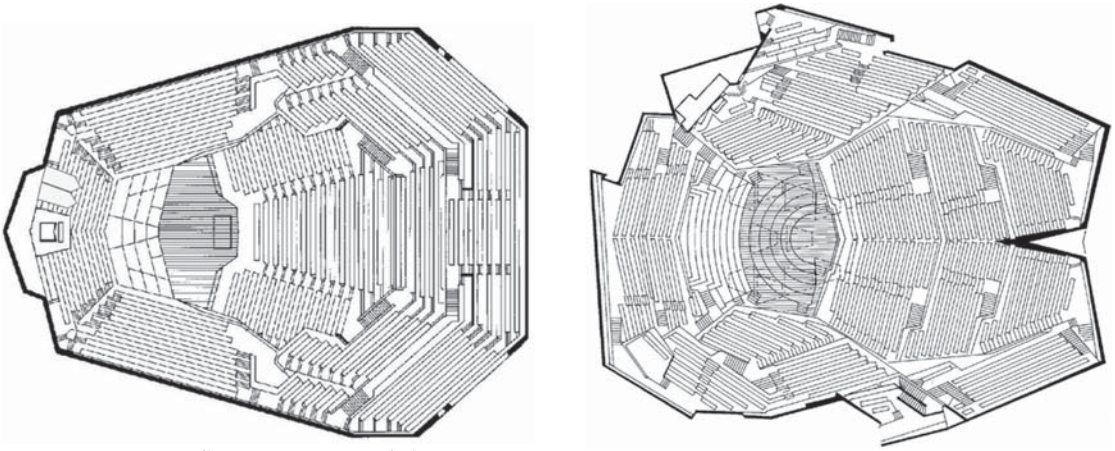
Figure 7-28. Examples of acoustically favorable ceiling designs.

to an adequately chosen geometry and size of the surfaces involved in this reverberation-time generating mechanism, it is possible to reduce the reverberation time in the low frequency range in a desirable fashion, while the sound impression is not deprived of its stimulating sound energy by sound absorption measures. In the concert hall of the Gewandhaus in Leipzig,[33] the room has its widest extension in the rear audience area, so that here, as a result of simulation measurements in a physical model (see Section 7-3), the height of the ceiling was chosen to have its maximum, Fig. 7-29A and B. In contrast to that, the maximum room width of the concert hall of the Philharmonie Berlin, Fig. 7-29C and D is in the region of the platform. To realize an optimum reverberation time, the maximum room height has to be above the platform.

This also explains why it was necessary to arrange room-height reducing panels in this concert hall. With music, the ceiling above the performance zone must neither fall below nor exceed a certain height in order to support the mutual hearing of the musicians and to avoid simultaneously the generation of disturbing reflections. According to reference 3, the lower limit of the ceiling height in musical performance rooms is 5 to 6 m (16 to 19 ft), the upper limit about 13 m (43 ft).

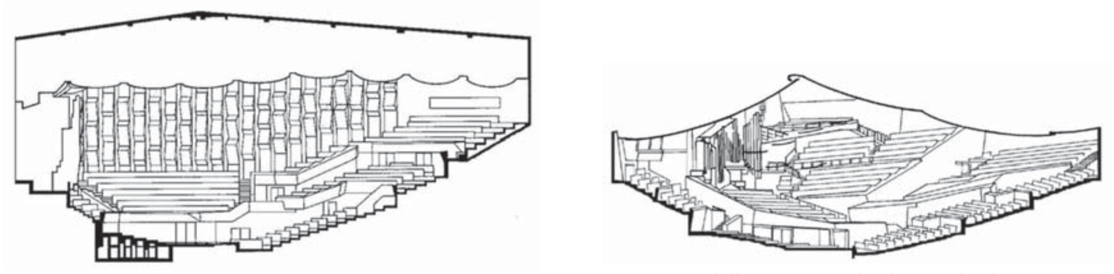
In large rooms for concert performances, the ceiling configuration should provide clarity-enhancing sound reflections in the middle and rear audience areas and simultaneously avoid disturbing reflections via remote boundary surfaces. Owing to the geometrical reflection, a plain ceiling arrangement, Fig. 7-30A, supplies only a slight portion of sound energy to the rear reception area, but in the front area (strong direct sound), the sound reflected by the ceiling is not required. In the rear ceiling area, however, the sound energy is reflected towards the rear wall from where it is returned, according to the unfavorable room geometry, as a disturbing echo (so-called theater echo), to the talker or the first listener rows. Keeping this in mind, the ceiling surfaces above the performance zone and in front of the rear wall should point perpendicular towards the middle seating area, Figs. 7-30B to D.

Monotonically curved ceilings in the shape of barrel vaults or cupolas show focusing effects, which in the neighborhood of such focuses may produce considerable disturbances in the listener or performer areas. The center



A. Gewandhaus Leipzig, ground plan.

C. Philharmonie Berlin, ground plan.



B. Gewandhaus Leipzig, longitudinal section.

D. Philharmonie Berlin, longitudinal section.

Figure 7-29. Various concert halls.

of curvature should therefore be above half the total height of the room or below twice the height, Fig. 7-31.

$$r \le \frac{h}{2} \text{ or } r \ge 2h \qquad (7\text{-}48)$$

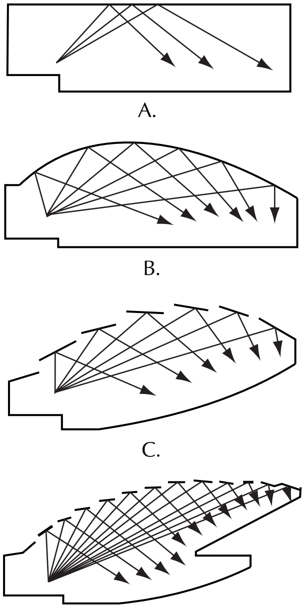


Figure 7-30. Ceiling configurations for obtaining energy-rich initial reflections in the middle and rear listener areas.

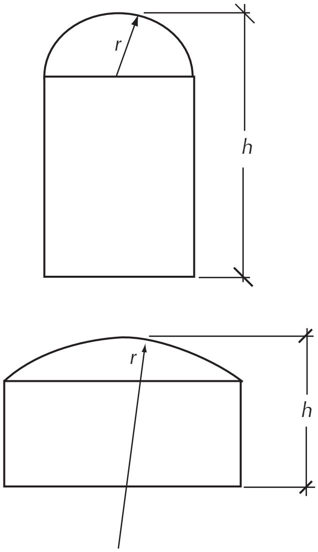


Figure 7-31. Focusing due to vaulted ceilings.

According to Cremer[9] it is guaranteed in this case that there do not originate from the curved ceiling any stronger reflections towards the receiving area than from a plain ceiling at apex height.

7.3.3.2.3 Balconies, Galleries, Circles

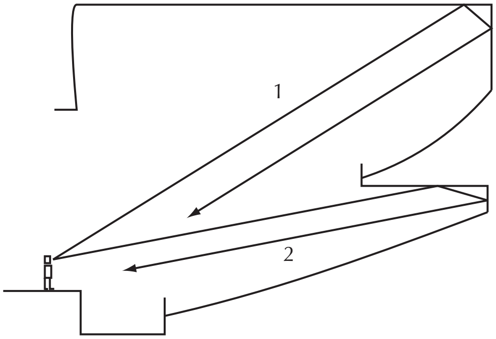
With proper arrangement and dimensioning, balconies and circles may have an acoustically favorable effect, since they contribute to a broadband diffuse sound dispersion and are also able to supply initial reflections for enhancing clarity and spatial impression. In this respect it is necessary, however, to decide if these reflections are desirable. Fig. 7-32A shows a graph of long-delayed sound reflections which give rise to very disturbing echo phenomena, the so-called theater echoes. Instrumental in the generation of these reflections is, first of all, the rear wall in combination with horizontal architectural elements (circle, gallery, balcony, ceiling). The disturbing effect of these sound reflections has to be avoided. Protruding balconies are, thanks to their horizontal depth, capable of shading these corner reflectors and to turn the reflections into useful sound, Fig. 7-32B.

The arrangement of far protruding circles is acoustically problematical with regard to the depth D (distance of the balustrade or of a room corner above it from the rear wall) and the clearance height H of the circle above the parquet or between two circles arranged one above the other. If the protrusion is very deep, the room area situated below it is shaded against reverberant sound and clarity-enhancing ceiling reflections. This area may be cut off from the main room and have an acoustic pattern of its own with strongly reduced loudness, unless certain construction parameters are observed, Fig. 7-33.[2,7,9]

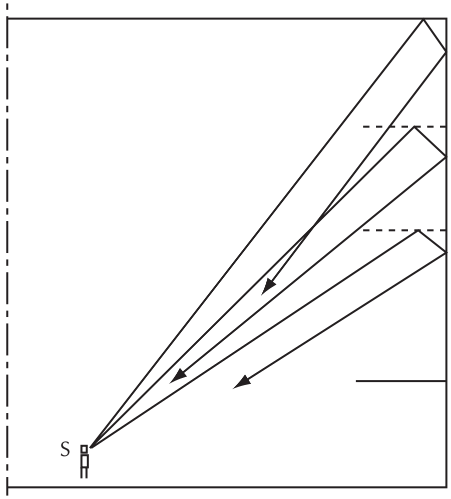
7.3.3.3 Room Topography

7.3.3.3.1 Sloping of Tiers, Sight Lines

For all room-acoustical parameters describing time and registers clarity, the energy proportion of direct sound and initial reflections is of great importance. With the sound propagating in a grazing fashion over a plain audience area there occurs a strong, frequency-dependent attenuation (see Section 7.3.4.4.4). Also, visually, such a situation implies considerable disadvantages by obstruction of the view towards the performance area. These disturbing effects are avoided by a sufficient and, if possible, constant superelevation of the visual line.

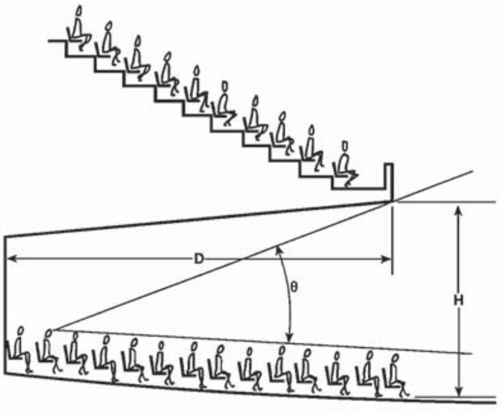


A. So-called theater echo.

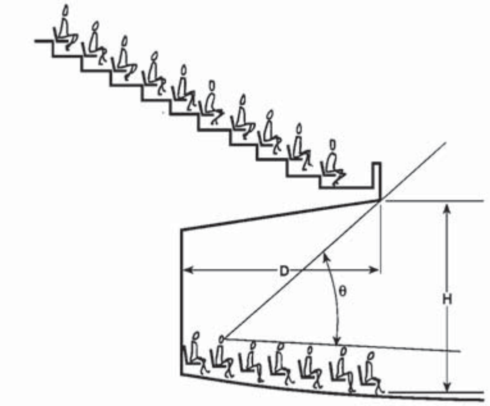


B. Edge reflections under circles and galleries.

Figure 7-32. Echo phenomena due to edge reflections.



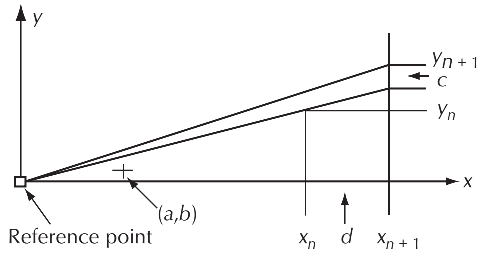
A. Music and opera houses, multigenre theaters.



B. Concert halls.

Room	Music and opera houses, multigenre theaters, Fig. 6-33A	Concert hall Fig. 6-33B
Circle depth D	$2H$	$< H$
Angle θ	25°	45°

Figure 7-33. Geometry of circle arrangement in A. Music and opera houses, multigenre theaters and B. concert halls.



(a,b) = Coordinates of the first row (eye level).
(xn, yn)) = Coordinates of the nth row.
(xn + 1 , yn + 1) = Coordinates of the (n + 1)th row.
c = Sight line superelevation (requirement: c = constant).
d = Tier spacing.
y = Eye level = tier level + 1.2 m.

Figure 7-34. Sloping of tiers (schematic view).

According to Fig. 7-34, this is the superelevation of the visual line (virtual line between eye and reference point) of a tier $n + 1$ as against tier n.

With a tier arrangement having a constant step height (continuous sloping in the longitudinal direction of the room), it is not possible to achieve a constant superelevation c. Mathematically it is the curve of a logarithmic spiral in which the superelevation increases alongside the distance from the reference point that realizes a constant superelevation of the visual line.[34]

As this implies, however, steps of different height for the individual tiers, a compromise must be found by either adapting the step height or by combining several tiers in small areas of constant sloping. In concert halls, the areas arranged in the shape of vineyard terraces (see Section 7.3.3.2.1) constitute, in this respect, an acoustically and optically satisfactory solution.

The eye level $y(x)$ is calculated with

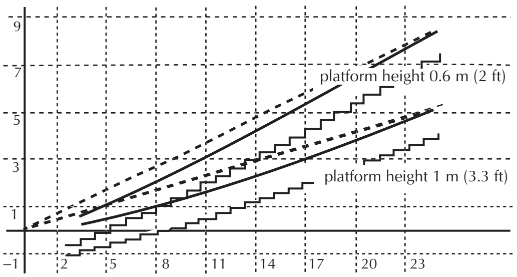
$$y = y_0 + \frac{cx}{d}\ln\frac{x}{a} + \frac{x}{a}(b - y_0) \tag{7-49}$$

where,

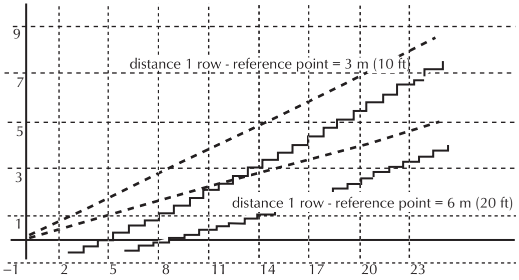
for $y_0 = 0$: $y = \frac{cx}{d}\ln\frac{x}{a} + \frac{bx}{a}$

The superelevation of the visual line should amount to at least 6 cm (2.5 in).

For estimating the required basic sloping of tiers one should keep in mind that the platform must be completely observable from all seats during the performances. The reference point to be chosen to this effect should, if possible, be the front edge of the platform. Using a reasonable platform height between 0.6 to 1 m (2 to 3.3 ft), the results of the sloping values are shown in Fig. 7-35A.



A. Sloping of tiers for different platform heights of 0.6 m and 1 m and constant distance of 3 m between first row and reference point.



B. Sloping of tiers for constant platform height of 0.6 m and different distance of 3 m and 6 m between first row and reference point.

Tier – Steps, eye line – solid, sight line – dashed.

Figure 7-35. Effect of sloping tiers.

By increasing the distance between first tier and viewing point (observable area of the platform) it is, of course, possible to notably reduce the necessary sloping of tiers, Fig. 7-35B.

With a plain parquet arrangement, which is the case in concert halls serving also for banquets or with classicistic architecture (Musikvereinssaal Vienna, Konzerthaus Berlin, Symphony Hall Boston, Herkulessaal Munich, etc.), a certain though normally somewhat dissatisfactory compensation is possible by means of an appropriate vertical staggering within the performance area (especially feasible for concert performances). For a basic platform height of 0.6 to 0.8 m (2 to 2.6 ft) it is possible to derive the theoretical required sloping heights from Eq. 7-49. With a length of a plain seating area of about 14 m (46 ft), a vertical staggering of the musicians on the platform must be about 3 m (10 ft), and with 18 m (59 ft) vertical staggering must be 4 m (13 ft). These values are generally not easy to realize, but show the necessity of an ample vertical staggering of the orchestra on the platform in rooms with plain parquet arrangement. However, if the optimum sloping of tiers according to Eq. 7-49 is realized on principle, it is possible for a sound source situated in the middle of the orchestra about 6 m (20 ft) behind the front edge of the platform to achieve the required view field angle by an elevation of 0.25 m (≈1 ft), and for the entire depth of the orchestra arrangement this elevation amounts to only about 1 m (≈3.3 ft). In concert halls with a sufficient sloping of tiers in the audience area, vertical staggering of the orchestra plays no more than a subordinate role for the unhindered direct sound supply to the audience area.

7.3.3.3.2 Platform Configuration in Concert Halls

With concert performances, the performance area for the orchestra (platform) must be an acoustical component of the auditorium, which means that both sections of the room must form a mutually attuned unity. This unity must not be disturbed by intermediate or other built-in elements. Any individual room-acoustical behavior of its own of a too small concert stage enclosure must be avoided. As used to be the case with many opera houses, this has a sound coloration deviating from that of the main auditorium and will be perceived as alienated. The volume of a concert stage enclosure should be at least 1000 m^3 (35,300 ft^3).[36] The sloping angles of the lateral boundary walls, referred to the longitudinal axis of the room, should be relatively flat. Takaku[36] defines an inclination index K according to Eq 7-50

$$K = \frac{\sqrt{\frac{WH}{\pi}} - \sqrt{\frac{wh}{\pi}}}{D} \tag{7-50}$$

where,

K is the inclination index,

W is the proscenium width,

H is the proscenium height,

w is the width of the rear wall,
h is the height of the rear wall,
D is the enclosure depth.

Optimum conditions for the mutual hearing of the musicians are achieved for a concert stage enclosure in the shape of a truncated pyramid, Fig. 7-36, with $K \leq 0.3$.[36]

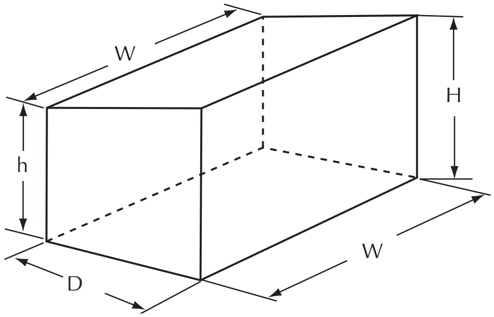


Figure 7-36. Geometrical parameters of a concert enclosure.

The more pronounced the diffuse subdivision of the inner surfaces of the concert stage enclosure, the smaller results the dependence of the room-acoustical parameters on the inclination index *K*.

If the platform boundaries are not formed by acoustically favorable solid wall and ceiling surfaces, additional elements have to be installed. The surface-related mass of the planking of these platform boundary surfaces should be chosen in such a way that the sound energy reduction by absorption results as little as possible. (The thinner the boundary walls the higher the low-frequency absorption.) To this effect, area-related masses of about 20 kg/m^2 (0.85 lbs/ft^2) are generally sufficient, in the neighborhood of bass instruments about 40 kg/m^2 (1.7 lbs/ft^2).

The vibration ability of the platform floor has only an insignificant influence on its sound radiation. With a relatively thin platform floor (12.5 mm (0.5 in) plywood[37] there may well result a sound amplification of between 3 dB to 5 dB in the lower-frequency range, but one should also not forget in this respect the positive psychological feedback of a vibrating floor on the players.[3] As a rule, the area-related mass of the platform floor should not fall below 40 kg/m^2.

By comparison with a rigid floor, a vibrating platform floor has, for the sound radiation of the bass string instruments with pizzicato play (faster decay resulting in a dry sound), the disadvantage of a reduced airborne sound energy, which can, however, be technically compensated with bow strokes.[3] This is why the platform floor should be frequency-tuned as low as possible.

The platform boundary surfaces should be structured in such a way that the mutual hearing of the musicians is supported, disturbing echo phenomena (e.g., by parallel wall surfaces) are avoided, and a well-mixed sound pattern gets radiated into the audience area. Obtaining a thorough mixing of the sound pattern requires a frequency-independent substructure of the boundary surfaces.

The space required per musician is about 1.4 m^2 (15 ft^2) for high-pitched string and brass instruments, 1.7 m^2 (18 ft^2) for low-pitched string instruments, 1.2 m^2 (13 ft^2) for woodwind instruments and 2.5 m^2 (27 ft^2) for the percussion. From this one can infer that, with due consideration of the participation of soloists (tail piano, etc.), the area of a concert platform (without choir) should generally not fall much below 200 m^2 (2200 ft^2), in which case the width should be about 18 m (60 ft) at the level of the high-pitched strings, and the maximum depth about 11 m (36 ft).

Depending on the sloping of tiers in the audience area (see Section 7.3.3.3.1), a vertical staggering of the orchestra is necessary especially if the audience area in the parquet is level or only slightly sloping. In the Musikvereinssaal Vienna the level difference on the platform is 1.8 m (6 ft), in the Berliner Philharmonie, which was destroyed during WWII, it was 2.8 m (9.2 ft). In such a case it is necessary to have one step in the string group approximately 250 mm (10 in), the following steps to and between the two rows of woodwind instruments should each be 500 mm (20 in) high. For the brass instruments or the percussion a further step of about 150 mm (6 in) is sufficient.

A choir, which in the staging of a grand concert, is normally lined up behind the orchestra, can profit only from the lateral wall surfaces and the ceiling of the room with regard to clarity-enhancing sound reflections, the floor area is shaded. Since according to Meyer[3] the main radiation axis of the singers' strongest sound fractions is inclined about 20° downwards, the choir line-up should be relatively steeply staggered in order to insure clarity and definition of articulation in the choir sound. With a flat line-up, however, only reverberance is increased. This is perceived as disturbing in rooms with a long reverberation time, whereas it may be rather desirable in reverberation-poor rooms. The optimum value of vertical staggering within a choir is about 45°—i.e., the steps should be equal in breadth and height in order to enable simultaneously an unhindered sound radiation to the lateral boundary surfaces of the room.[3]

7.3.3.3.3 Orchestra Pit

On principle, the arrangement of the orchestra with musical stage plays in the so-called pit at the border line between stage and auditorium is acoustically unfavorable by comparison with orchestra arrangements on the stage (e.g., stage music), but has developed historically from the performing practice in the 19th century. In most baroque theaters the musicians were seated either at the same level as the first listeners' rows or only a few steps lower.[3] They were separated from the audience area only by an about 1 m (3.3 ft) high balustrade. With the introduction of an orchestra pit, the visual contact between listener and stage was later reduced, especially when the orchestras grew larger. Room-acoustical shortcomings lie herewith in the problem of balance between singing/speech on stage and the accompanying orchestra in the pit. Owing to the size and equipment of the stage area, the loudness of the singers gets altered with growing distance from the orchestra so that balance problems increase especially in case of low singing loudness and unfavorable pitch levels.

A further aspect concerns the register and time correspondence between stage and pit on which depend the intonation and ensemble playing.

The geometrical separation between the two performance areas (stage and orchestra pit) should, in modern opera houses, be as little as possible, not only in dependence on dramaturgical arguments, but also for visual and functional reasons. Consequently, the orchestra pit slides beneath the stage, so as to avoid that the distance of the first rows from the stage increases still further. The practicability of the thus formed covered area of the orchestra pit (proscenium area), required for dramaturgical reasons, implies that the covered area becomes bigger and bigger, while the open coupling space of the pit to the auditorium gets smaller and smaller. The orchestra pit thus becomes an independent room tightly packed with musicians and with low boundary surfaces, a low-volume index, and a nonreflecting subceiling (opening) representing the outlet for irradiation of the more or less well mixed orchestra sound to the auditorium. Owing to the reduced distance of the musicians from the boundary surfaces, the sound pressure level in the pit increases by up to about 4 dB, whereby the mutual hearing of the musicians is supported for low-volume playing. With increased loudness the mutual hearing gets disadvantageously limited to loud instrumental groups in the low- and medium-frequency range.

Sound-absorbing wall or ceiling coverings or adjustable wall elements with preferential effect in the low- and medium-frequency range, arranged in the neighborhood of loud instruments, reduce the loudness desirably, but not the direct sound irradiation into the auditorium. This supports the clarity of the sound pattern.[3] If the orchestra pit level is very low, about 2.5 m (8.2 ft), direct sound fractions reach the parquet level only by diffraction, causing the sound pattern to be very bass-accentuated. Brilliance and temporal clarity become adequate only at those places where visual contact to the instrument groups is given (circles).

Acoustic improvement of this situation may be achieved on the one hand by a wider opening of the orchestra pit, so that energy-rich initial reflections are enabled via a corresponding structure of the adjacent proscenium area (proscenium ceiling and side wall design). On the other hand the pit depth should not exceed certain limits. By means of subjective investigations with varying height of the pit floor, optimum solutions may easily be found here in combination with an adequate positioning of the instruments in the pit. With a balustrade height of about 0.8 m (2.6 ft), lowering the front seating area of the pit floor (high-pitch strings) to about 1.4 m (4.6 ft) produces generally good acoustical conditions. Towards the rear the staggering should go deeper.

Provided the orchestra plays with adapted loudness, an acceptable solution consists in an almost complete opening of the orchestra pit towards the proscenium side walls and an as little as possible covering towards the stage. If the open area amounts to at least 80% of the pit area, the orchestra pit becomes acoustically part of the auditorium and the unity of the sound source is insured also with respect to coloration (example: Semperoper Dresden). Another solution consists of an almost completely covered orchestra pit with a small coupling area to the auditorium. This requires, however, a correspondingly large pit volume with a room height of at least 3 m (10 ft) (example: Festspielhaus Bayreuth). Common opera houses lie with their orchestra pit problems half-way between these two extremes. If there is a large orchestration accommodated in the pit, less powerful singers on the stage may easily become acoustically eclipsed. More favorable conditions can be obtained in this case by means of a pit covering, provided a sufficient volume is given, or by positioning the orchestra on a lower pit floor level.

Apart from the sound reflecting and sound absorbing boundary surfaces arranged in the pit for supporting the mutual hearing and the intonation, the inner faces of the pit balustrade should point perpendicularly towards the stage (slight inclinations on the side of the balustrade). In this way the stage is better supplied with initial reflections from the pit, whereas the pit receives a first reflection from sound sources on the stage. Convex curves (in the vertical domain) combined with a sound-absorbing effect in the low-frequency range are, in this respect, especially advantageous for making the supporting effect register-independent on the one hand and brilliance enhancing on the other hand. The edge of the stage above the pit vis-à-vis the conductor should be conformed geometrically in such a way that additional initial reflections are directed to the audience area. The lateral configuration of the pit opening, combined with an appropriate subconstruction of the proscenium side wall, should insure a maximum of sound reflections towards the pit and the stage.

7.3.4 Secondary Structure of Rooms

7.3.4.1 Sound Reflections at Smooth Plane Surfaces

With the reflection of sound rays from boundary surfaces, one can principally define three types of reflection which differ from one another by the relation between the linear dimensions and the wavelength and by the relation between the reflected and the incident sound ray, Fig. 7-37.

- **Geometrical reflection**, Fig. 7-37A: $b < \lambda$, $\alpha = ß$ (specular reflection according to the reflection law in one plane perpendicular to the carrier wall).
- **Directed (local) reflection**, Fig. 7-37B: $b > \lambda$, $\alpha = ß$ (specular reflection according to the reflection law, referred to the effective structural surface).
- **Diffuse reflection**, Fig. 7-37C: $b \approx \lambda$, (no specular reflection, without a preferred direction).

A geometrical sound reflection occurs at a sufficiently large surface analogously to the reflection law of optics: the angle of incidence α is equal to the angle of reflection ß and lies in a plane perpendicular to the surface, Fig. 7-38. This reflection takes place only down to a lower limit frequency f_{low}

$$f_{low} = \frac{2c}{(b\cos\alpha)^2} \times \frac{a_1 a_2}{a_1 + a_2} \qquad (7\text{-}51)$$

where,
c is the velocity of sound in air.

Below f_{low} the sound pressure level decay amounts to 6 dB/octave.[38]

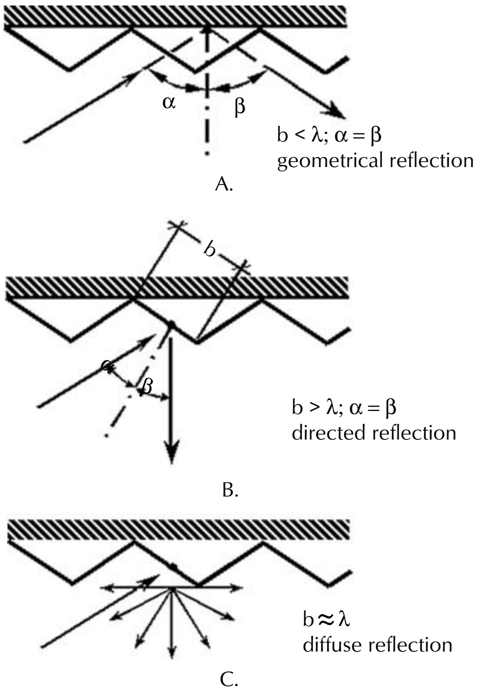


Figure 7-37. Basic sound reflections at smooth, plane surfaces.

Eq. 7-51 has been graphically processed, Fig. 7-39.[3] With a reflector extension of 2 m (6.6 ft) at a distance of 10 m (3 ft) each from the sound source and to the listener, the lower limiting frequency is, for example, about 80 Hz with vertical sound incidence and about 1600 Hz with an incidence angle of 45°. If this reflector is installed as a panel element in the front part of the platform, the frequency region of the sound reflections is about one octave lower with almost platform-parallel arrangement than in a 45° inclined position. The desired limiting frequency goes down to lower values under the following circumstances:

- The bigger the effective surface.
- The nearer to the sound source and to the listener the reflector is installed.
- The smaller the sound incidence angle.

Apart from the geometry of the reflectors, the area-related mass of the same also has to be consistent with certain limit values in order to obtain a reflection with as little a loss as possible. If the reflectors are employed for speech and singing in the medium and high-frequency ranges, a mass of about 10 kg/m^2 (1.7 lbs/ft^2) is sufficient (e.g. a 12 mm (½ in) plywood plate). If the effective frequency range is expanded to bass instruments, a mass of about 40 kg/m^2 (1.7 lbs/ft^2) has to be aspired (e.g., 36 mm [1.5 in] chipboard). With reflectors additionally suspended above the performance zone, the statically admissible load often plays a

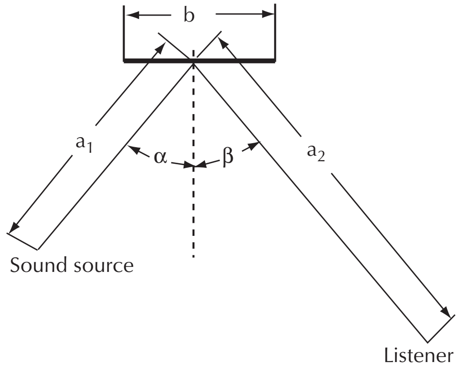


Figure 7-38. Geometrical sound reflections.

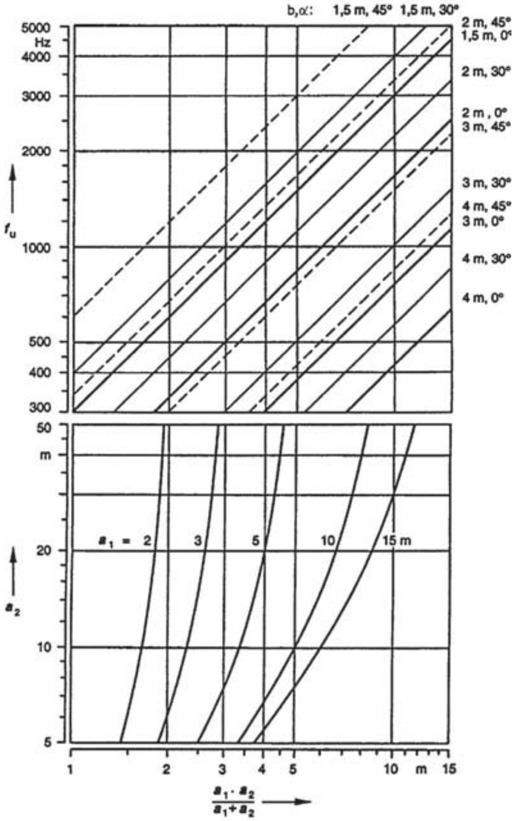


Figure 7-39. Minimum size of surfaces for geometrical sound reflections.

restrictive role for the possible mass of the reflectors. For spoken performances an area-related mass of 5 to 7 kg/m² (0.2 to 0.3 lbs/ft²) may still produce acceptable results, to which effect plastic mats of high surface density are suitable. The additional room-acoustical measure usually employed for enhancing the sound reflection of bass instruments with music performances consists in appropriate wall surfaces, so that the installation of heavy panels can be abandoned. In this case an area-related mass of 20 kg/m² (0.8 lbs/ft²) is sufficient.

If a multiple reflection occurs close to edges of surfaces, there results, if the edge is at right angle to the surface, a sound reflection with a path parallel to the direction of the sound incidence, Fig. 7-40. In corners, this effect acquires a 3D nature, so that the sound always gets reflected to its source, independently of the angle of incidence. With long travel paths it is possible that very disturbing sound reflections are caused at built-in doors, lighting stations, setoffs in wall paneling, which for the primary structure of a room are known as "theater echo" (see Section 7.3.3.2.2).

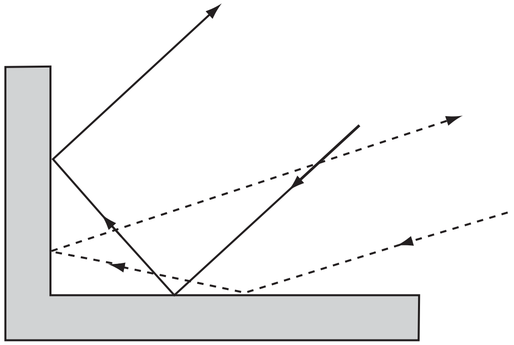


Figure 7-40. Multiple reflection in room edges.

7.3.4.2 Sound Reflection at Smooth Curved Surfaces

If the linear dimensions of smooth curved surfaces are much bigger than the wavelength of the effective sound components, the sound is reflected from these surfaces according to the laws of concentrating reflectors. Concavely curved 2D or 3D surface elements may, under certain geometrical conditions, lead to sound concentrations while convex curvatures always have a sound scattering effect.

For axis-near reflection areas (incident angle less than 45°) of a surface curved around the center of curvature *M,* it is possible to derive the following important reflection variants, Fig. 7-41.

Circular Effect. The sound source is located in the center of curvature *M* of the reflecting surface, Fig. 7-41A. All irradiated sound rays become concentrated in *M* after having covered the radius twice, so that a speaker may, for instance, be heavily disturbed by his own speech.

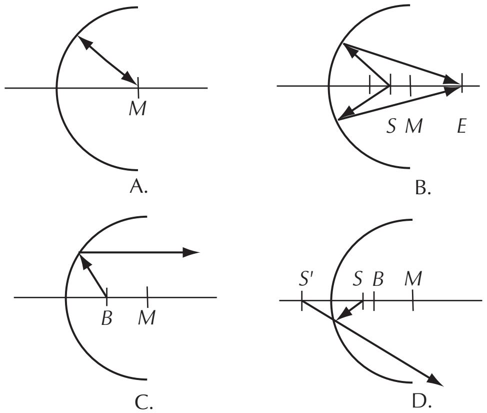


Figure 7-41. Sound reflection at smooth, curved surfaces.

Elliptical Effect. If the sound source is located between half the radius of curvature and the full radius of curvature in front of the reflecting surface, a second sound concentration point is formed outside the center of curvature, Fig. 7-41B. If this second focus is located within the performance zone or the audience area, it is perceived as very disturbing, since distribution of the reflected sound is very unbalanced. With extended sound sources like an orchestra, curved surfaces of this kind produce a heavily register-dependent sound balance.

Parabolic Effect. If in a rather narrow arrangement the sound source is located at half the center of curvature, Fig. 7-41C, the curved surface acts like a so-called parabolic reflector that generates an axis-parallel bundle of rays. This produces, on the one hand, a very uniform distribution of the reflected portion of the sound irradiated by the source, but on the other hand there occurs an unwanted concentration of noise from the audience area at the location of the sound source.

Hyperbolic Effect. If the distance of the sound source from the curved surface is smaller than half the radius of curvature, Fig. 7-41D, the reflecting sound rays leave the surface in a divergent fashion. But the divergence is less and thus the sound intensity at the listener's seat is higher than with reflections from a plain surface.[2] The acoustically favorable scattering effect thus produced is comparable to that of a convexly curved surface, but the diverging effect is independent of the distance from the curved reflecting surface.

7.3.4.3 Sound Reflections at Uneven Surfaces

Uneven surfaces serve as the secondary structure of directional or diffuse sound reflections. This refers to structured surfaces with different geometrical intersections in the horizontal and vertical planes (rectangles, triangles, sawtooth, circle segments, polygons) as well as 3D structures of geometrical layout (sphere segments, paraboloids, cones, etc.) and free forms (relievos, moldings, coves, caps, ornaments, etc.). Also by means of a sequence of varying wall impedances (alternation of sound reflecting and sound absorbing surfaces), it is possible to achieve a secondary structure with scattering effect.

To characterize this sound dispersion of the secondary structure one makes a distinction between a degree of diffusivity d and a scattering coefficient s.

Typically for the homogeneity of the distribution of the sound reflections is the so-called frequency-dependent degree of diffusivity d.[47]

$$d = \frac{\left(\sum_{i=1}^{n} 10^{\frac{L_i}{10}}\right)^2 - \sum_{i=1}^{n}\left(10^{\frac{L_i}{10}}\right)^2}{(n-1)\sum_{i=1}^{n}\left(10^{\frac{L_i}{10}}\right)^2} \tag{7-52}$$

This way angle-dependent diffusion balloons may be generated. Depending on the number of n receiver positions hi-res=level values are supplied to form the balloon.

High diffusion degrees close to one will be reached for half-cylinder or half-sphere structures. Nevertheless the diffusion degree d is more or less a qualitative measure to evaluate the homogeneity of scattering.

On the other side and as a quantitative measure to characterize the amount of scattered energy in contrast to the specular reflected or absorbed energy, the frequency-dependent scattering coefficient s is used.[50]

This scattering coefficient s is used in computer programs to simulate the scattered part of energy especially by using ray tracing methods.

The coefficient s will be determined as the ratio of the nonspecular (i.e., of the diffuse reflected) to the overall reflected energy.

$$\begin{aligned} s &= \frac{diffuse - reflected - Energy}{overall - reflected - Energy} \\ &= \left(1 - \frac{geometric - reflected - Energy}{overall - reflected - Energy}\right) \end{aligned} \tag{7-53}$$

The measurement and calculation of the scattering coefficient under random sound impact take place in the reverberation chamber.[39,48,49]

All these parameters don't say too much about the angular distribution of the reflected sound energy. But there exist many examples of rooms in which the secondary structure is intended to realize a directional reflection in which the angle of sound reflection does not correspond to the angle of sound incidence, as referred to the basic surface underlying the primary structure. In this case of directional sound reflection, one has to consider parameters determining, among other things, the diffusivity ratio D_{diff} and the maximum displacement d_α, Fig. 7.42.[42]

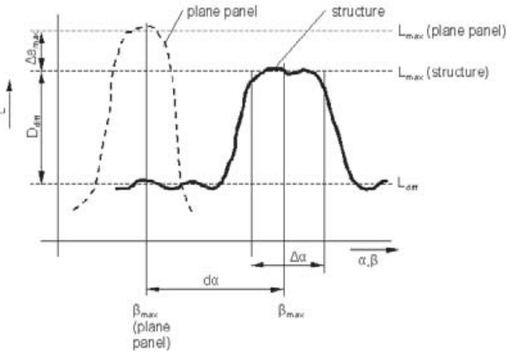


Figure 7-42. Parameters for characterizing the directivity of uneven surfaces.

- *Diffusivity ratio* D_{diff}: sound pressure level difference between the directional and the diffuse sound components L_{max} and L_{diff}, respectively.
  → Characterizes the directional effect of a structure.
- *Attenuation of Maximum* Δa_{max}: sound pressure level difference between the directional reflection (local maximum, β_{max}) of the structure, as compared with a plain surface.
  → Characterizes the sound pressure level of the reflection.
- *Displacement of Maximum* $d\alpha$: angle between geometrical and directional reflections.
  → Characterizes the desired change of direction of the reflection.
- Angular range of uniform irradiation $\Delta\alpha$: 3 dB bandwidth of the reflection.
  → Characterizes the solid-angle range of uniform sound reflection.

Guide values D_{diff} for the octave midband frequency of 1000 Hz are based on subjective sound-field investigations in a synthetic sound field, Table 7-9.

Table 7-9. Perception of Diffuse and Directed Sound Reflections as a Function of the Diffusivity Ratio D_{diff}

Perception	D_{diff} in dB
Ideal diffuse sound reflection	0
Diffuse sound reflection	< 3
Range of appropriate perception of diffuse and directed sound reflections	3–10
RT around 1.0 s with energy-rich ceiling reflections	2–6
RT around 2.0 s with energy-rich ceiling reflections	4– 8
Spatial sound fields with low direct sound energy, but big part of lateral reflections	6– 8
Sound fields with high direct sound energy—e.g., more distant listener groups	3– 6
Low sound energy of ceiling reflections and big part of lateral sound	8–10
Directed sound reflection	>10
Ideal directed sound reflection	∞

An example for a sawtooth structure is shown in Fig. 7-43. This side wall structure has at a sound incidence angle of 50° and a speech center frequency of about 1000 Hz, energy-rich directional sound reflections ($D_{diff} \geq 10$ dB) with a displacement of maximum of $d\alpha = -20°$ (reflection angle 30°). Additionally directional and diffuse sound components being perceptible from about 3000 Hz ($D_{diff} = 6$ to 8 dB). The attenuation of maximum Δa_{max} was 5 dB at 1000 Hz and 11 dB at 5000 Hz by comparison with a carrier panel of geometrical sound reflection.

Periodical structures of elements having a regular geometrical cut (rectangle, isosceles triangle, sawtooth, cylinder segment) may show high degrees of scattering, if the following dimensions are complied with, Fig. 7-44.[3,40]

For a diffuse scattering in the maximum of the speech-frequency range, the structure periods are therefore about 0.6 m (2 ft), the structure widths between 0.1 to 0.4 m (0.33 to 1.3 ft), and the structure heights maximally about 0.3 m (1 ft). With rectangular structures the sound scattering effect is limited to the relatively small band of about one octave, with triangular structures to maximally two octaves. Cylinder segments or geometrical combinations can favorably be used for more broadband structures, Fig. 7-43. In a wide-frequency range between 500 Hz and 2000 Hz, a cylinder segment structure is sufficiently diffuse, if the structure width of about 1.2 m (4 ft) is equal to the structure period, and the structure height is between 0.15 and 0.20 m (0.5 and 0.7 ft). With a given structure height h and a given structure width b it is, according to Eq. 7-54, possible to calculate the required curvature radius r as

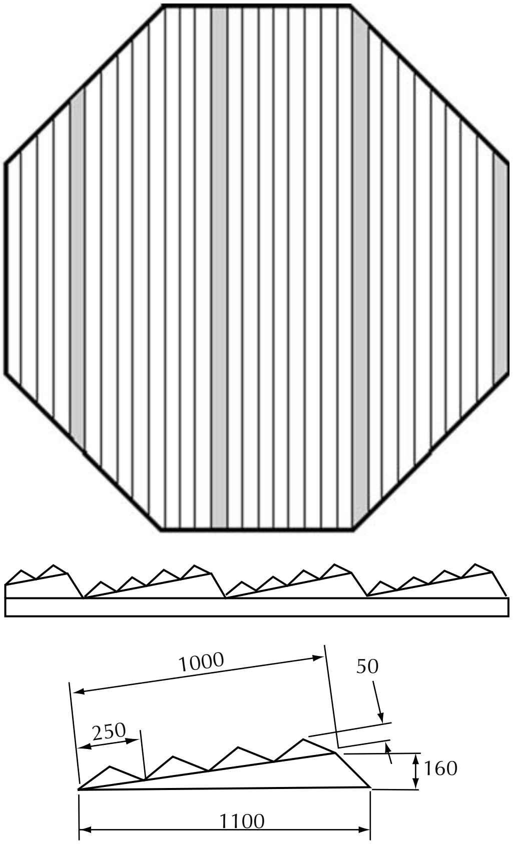
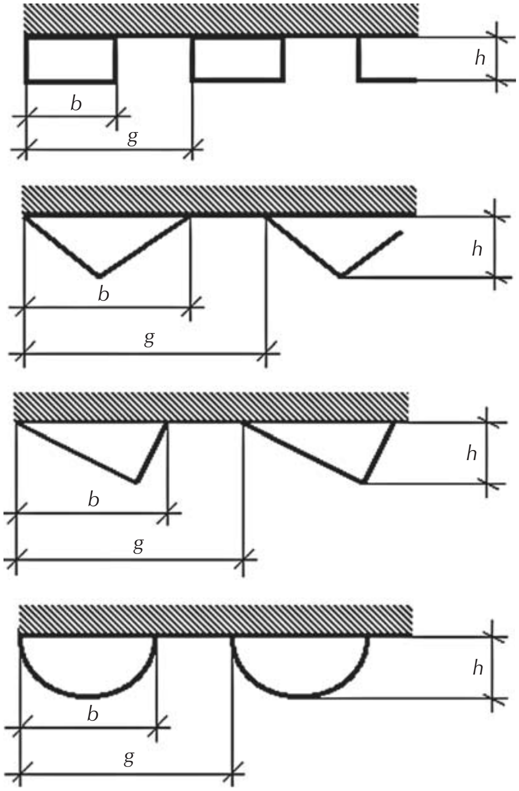


Figure 7-43. Example of an acoustically active sawtooth structure. Measures in mm.

$$r = \frac{\left(\frac{b}{2}\right)^2 + h^2}{2h} \tag{7-54}$$



Structure	Structure period *g*	Structure width *b*	Structure height *h*
Rectangle	≈(1–2)λ	≈0.2g	≈0.2g
Isosceles triangle	≈(1–2)λ	≈(0.5g–0.67)g	≈(0.25–0.33)g
Sawtooth	≈2λ		≈0.33λ
Cylinder segment	≈(1–2)λ	≈(0.17–1.0)g	≈(0.25–0.5)g

Figure 7-44. Geometrical parameters at structures with rectangular, triangular, sawtooth-formed, and cylinder-segment-formed intersection.

A special form of a diffusely reflecting surface can be realized by lining up phase-grating structures of varying depths. Based on the effect of coupled λ/2 runtime units, these structures produce on the surface a local distribution of the reflection factor and hence of the sound particle velocity. Every component of this velocity distribution produces thereby a sound irradiation into another direction. If according to Schroeder[41] one distributes these reflection factors in accordance with the maximum sequences of the number theory (e.g., Barker code, primitive root diffusor PRD, square-law residual series QRD), and separates these trough structures from each other by thin wall surfaces, one obtains diffuse structures of a relatively broadband effect (up to two and more octaves), Fig. 7-45. With perpendicular sound incidence, the minimum frequency limit f_{low} for the occurrence of additional reflection directions is approximately

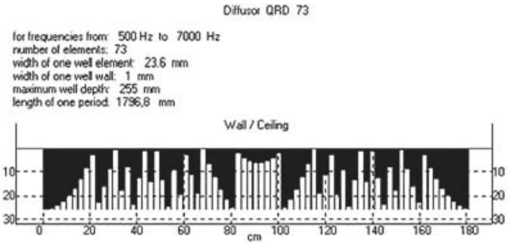
$$f_{low} \approx \frac{c}{2d_{max}} \tag{7-55}$$

where,

c is the velocity of sound in air in m/s (ft/s),

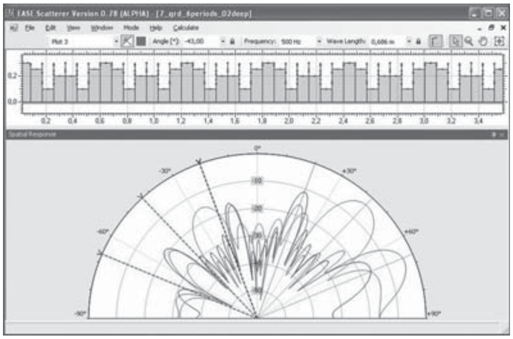
d_{max} is the maximum depth of structure in m (ft)

Nowadays calculation programs are available to calculate the scattering coefficients for angle-dependent sound impact by using Boundary Element Methods, Fig. 7-46.


Diffusor QRD 73

for frequencies from: 500 Hz to 7000 Hz
number of elements: 73
width of one well element: 23.6 mm
width of one well wall: 1 mm
maximum well depth: 255 mm
length of one period: 1796,8 mm

Figure 7-45. Schroeder diffusor with primitive root structure.



Angle [°]: -43,00
Frequency: 500 Hz
Wave Length: 0,686 m

Figure 7-46. Boundary Element Methods (BEM) based software tool for calculating scattering coefficients.

7.3.4.4 Sound Absorbers

Sound absorbers can occur in the shape of surfaces, built-in elements, pieces of furniture, or in the form of unavoidable environmental conditions (e.g., air) as well as arrangements conditioned by utilization of the room (e.g. spectators, decorations). According to their preferential effect in a determined frequency range one distinguishes on principle between

- Absorbers in the low-frequency range between approximately 32 Hz and 250 Hz.
- Absorbers in the medium-frequency range between approximately 315 Hz and 1000 Hz.
- Absorbers in the high-frequency range between approximately 1250 Hz and 12 kHz.
- Broadband absorbers.

For acoustical characterization of a sound absorber there serves its frequency-dependent sound absorption coefficient α or the equivalent sound absorption area A. For an area of size S one determines the equivalent sound absorption area A as

$$A = \alpha S \tag{7-56}$$

The sound power W_i being incident on an area of size S of a sound-absorbing material or a construction is designated as sound intensity I_i, part of which is reflected as sound intensity I_r, and the rest is absorbed as sound intensity I_{abs}. The absorbed sound intensity consists of the sound absorption by *dissipation* (transformation of the sound intensity I_δ in heat, internal losses by friction at the microstructure or in coupled resounding hollow spaces), and of the sound absorption by *transmission* (transmission of the sound intensity I_τ into the coupled room behind the sound absorber or into adjacent structural elements).

$$\begin{aligned} I_i &= I_r + I_{abs} \\ &= (I_r + I_\delta + I_\tau) \end{aligned} \tag{7-57}$$

With the *sound reflection coefficient* ρ defined as

$$\rho = \frac{I_r}{I_i} \tag{7-58}$$

and the *sound absorption coefficient* α as

$$\begin{aligned} \alpha &= \frac{I_\delta + I_\tau}{I_i} \\ &= \frac{I_{abs}}{I_i} \end{aligned} \tag{7-59}$$

as a sum of the *dissipation coefficient* δ

$$\delta = \frac{I_\delta}{I_i} \tag{7-60}$$

and the *transmission coefficient*

$$\tau = \frac{I_\tau}{I_i} \tag{7-61}$$

Eq. 7-57 becomes

$$\begin{aligned} 1 &= \rho + \delta + \tau \\ &= (\rho + \alpha) \end{aligned} \tag{7-62}$$

The transmission coefficient τ plays a role when considering the sound insulation of structural components. For nonmovable, monocoque, acoustically hard material surfaces (e.g., walls, windows), it is, according to Cremer,[48] possible to consider the frequency dependence of the transmission coefficient as a low-pass behavior which surpasses the given value up to a limit frequency f_τ. With a negligible dissipation coefficient it is furthermore possible to equate the transmission coefficient τ numerically to the sound absorption coefficient α.

7.3.4.4.1 Sound Absorption Through Porous Materials

The effect of sound absorption is based essentially on the transformation of sound energy in thermal energy by air particles moving in open, narrow, and deep pores. Closed pores like those existing in foamed materials used for thermal insulation are unsuited for sound insulation. For characterizing the materials the so-called *porosity* σ is used. This represents the ratio between open air volume V_{air} existing in the pores and the overall volume V_{tot} of the material

$$\sigma = \frac{V_{air}}{V_{tot}} \tag{7-63}$$

With a porosity of $\sigma = 0.125$, it is possible for high frequencies to obtain a maximum sound absorption coefficient of only $\alpha = 0.4$, and with $\sigma = 0.25$ of $\alpha = 0.65$. Materials with a porosity of $\sigma \geq 0.5$ enable a maximum sound absorption coefficient of at least 0.9. Usual mineral, organic, and naturally growing fibrous insulating materials feature porosities of between 0.9 and 1.0 and are thus very well suited for sound absorption purposes in the medium- and high-frequency ranges.[29]

Apart from porosity it is also the structure coefficient *s* and the flow resistance Ξ which influence the sound absorbing capacity of materials. The structure coefficient *s* can be calculated from the ratio between the total air volume V_{air} contained in the pores and the effective porous volume V_w

$$s = \frac{V_{air}}{V_w} \tag{7-64}$$

The insulating materials most frequently used in practice have structure factors of between 1 and 2—i.e., either the total porous volume is involved in sound transmission or the dead volume equals the effective volume. Materials with a structure factor of the order of ten show a sound absorption coefficient of maximally 0.8 for high frequencies.[8]

The flow resistance exerts an essentially higher influence on sound absorption by porous materials than the structure factor and the porosity. With equal porosity, for instance, narrow partial volumes offer a higher resistance to particle movement than wide ones. This is why the specific flow resistance R_s is defined as the ratio of the pressure difference before and behind the material with regard to the speed of the air flowing through the material v_{air}

$$R_s = \frac{\Delta p}{v_{air}} \tag{7-65}$$

where,

R_s is the specific flow resistance in Pa s/m (lb s/ft^3),

Δp is the pressure difference in Pa (lb/ft^2),

v_{air} is the velocity of the passing air in m/s (ft/s).

With increasing material thickness the specific flow resistance in the direction of flow increases as well.

7.3.4.4.2 Sound Absorption by Panel Resonances

Thin panels or foils (vibrating mass) can be arranged at a defined distance in front of a rigid wall so that the withdrawal of energy from the sound field in the region of the resonance frequency of this spring-mass vibrating system makes the system act as a sound absorber. The spring action is produced herewith by the rigidity of the air cushion and the flexural rigidity of the vibrating panel. The attenuation depends essentially on the loss factor of the panel material, but also on friction losses at the points of fixation.[43] The schematic diagram is shown in Fig. 7-47, where d_L is the thickness of the air cushion and m' the area-related mass of the vibrating panel.

The resonance frequency of the vibrating panel mounted in front of a rigid wall with attenuated air space and lateral coffering is calculated approximately as

$$f_R \approx \frac{60^*}{\sqrt{m'd_L}} \tag{7-66}$$

* 73 in U.S. units

where,

f_R is in Hz,

m' is in kg/m^2 (lb/ft^2),

d_L is in m (ft).

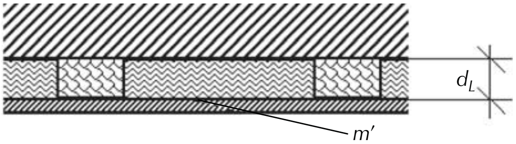


Figure 7-47. General structure of a panel resonator.

In practical design one should moreover take into account the following:

- The loss factor of the vibrating panel should be as high as possible.[43]

- The clear spacing of the coffering should be, in general, smaller in every direction than 0.5 times the wavelength in case of resonance, but not fall short of 0.5 m (1.7 ft).
- The minimum size of the vibrating panel must not fall short of 0.4 m^2 (4.3 ft^2).
- The air-space damping material should be attached to the solid wall so that the panel vibration is not impaired in any way.
- The sound absorption coefficient depends on the *Q* factor of the resonance circuit and amounts at the resonance frequency to between 0.4 and 0.7 with air-cushion damping and to between 0.3 and 0.4 without air-cushion damping. At an interval of one octave from the resonance frequency one must reckon that the sound absorption coefficient is halved.

An effective method for increasing the acoustically effective resonance frequencies of panel resonators consists of reducing the vibrating mass of heavy panels by means of holes arranged in defined patterns. In this case the correlations are governed by analogous regularities, if the area-related mass m' of the panels is replaced by the effective hole mass m'_L. For circular holes of radius R and a hole-surface ratio ε, Fig. 7-48, the hole mass is calculated as

$$m'_L = 1.2^{**}\frac{l^*}{\varepsilon} \qquad (7\text{-}67)$$

** 0.37 for U.S. system

where,

m'_L is the area-related air mass of circular openings in kg/m^2 (lb/ft^2),

l^* is the effective panel thickness with due consideration of the mouth correction of circular openings of radius R in meters (ft)

$$l^* \approx 1 + \frac{\pi}{2}R\,, \qquad (7\text{-}68)$$

ε is the hole-area ratio according to Fig. 7-47 for circular openings

$$\varepsilon = \frac{\pi R^2}{ab}\,. \qquad (7\text{-}69)$$

Provided the hole diameters are sufficiently narrow, the damping material layer arranged between the perforated panel and solid wall can be replaced by the friction losses produced in the openings. By using transparent materials—e.g., glass—it is possible to fabricate optically transparent, so-called micro-perforated absorbers. The diameters of the holes are in the region of 0.5 mm (0.02 in) with a panel thickness of 4 to 6 mm (0.16 to 0.24 in) and a hole-area ratio of 6%. For obtaining broadband sound absorbers, it is possible to resort to variable perforation parameters (e.g., scattered perforation) varying thickness of the air cushion and composite absorbers combined of various perforated panels.

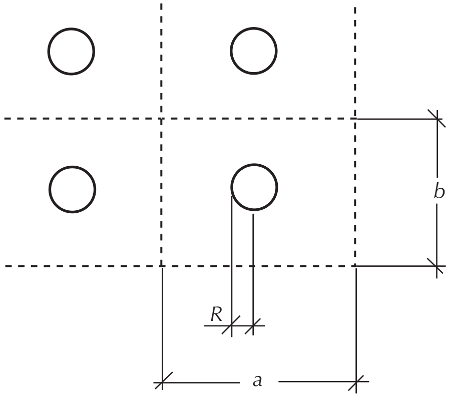


Figure 7-48. Hole-area ratio of perforated panels with round holes.

A very recent development are microperforated foils of less than 1 mm (0.04 in) thickness which also produce remarkable absorption when placed in front of solid surfaces. The transparent absorber foil can be advantageously arranged in front of windows either fixed or also as roll-type blinds in single or double layer.[44]

7.3.4.4.3 Helmholtz Resonators

Helmholtz resonators are mainly used for sound absorption in the low-frequency range. Their advantage, as compared with panel absorbers (see Section 7.3.4.4.2), lies in their posterior variability regarding resonance frequency and sound absorption coefficient as well as in the utilization of existing structural cavities which must not necessarily be clearly visible. According to Fig. 7-49, a Helmholtz resonator is a resonance-capable spring-mass system which consists of a resonator volume V acting as an acoustical spring and of the mass of the resonator throat characterized by the opening cross section S and the throat depth l. In resonance condition and if the characteristic impedance of the resonator matches that of the air, the ambient sound field is deprived of a large amount of energy. To this effect a damping material of a defined specific sound impedance is placed in the resonator throat or in the cavity volume.

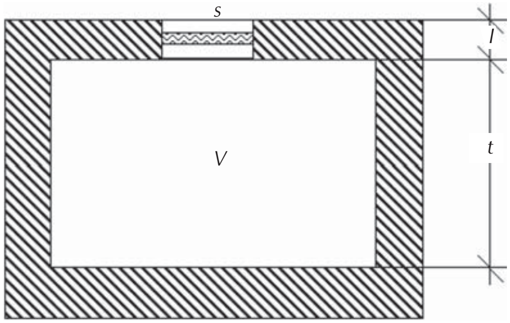


Figure 7-49. General structure of a Helmholtz resonator.

The resonance frequency of a Helmholtz resonator is generally calculated as

$$f_R = \frac{c}{2\pi}\sqrt{\frac{S}{V(l+2\Delta l)}} \tag{7-70}$$

where,
c is the speed of sound in air, approximately 343 m/s (1130 ft/s),
S is the cross-sectional area of the resonator in m^2 (ft^2),
V is the resonator volume in m^3 (ft^3),
l is the length of the resonator throat in m (ft),
$2\Delta l$ is the mouth correction.

In case of a square opening

$$2\Delta l \approx 0.9a\,,$$

where,
a is the edge length of the square opening.

For circular openings the resonance frequency f_R in Hz is calculated approximately from Eq. 7-70:

$$f_R \approx \frac{100^*R}{\sqrt{V(l+1.6R)}} \tag{7-71}$$

* 30.5 for U.S. units

where,
R is the radius of the circular opening in m (ft),
V is the resonator volume in m^3 (ft^3),
l is the length of the resonator throat in m (ft).

7.3.4.4.4 Sound Absorption by the Audience

The efficiency of sound absorption by the audience depends on many factors, for instance the occupation density, the spacing of seats and rows, the clothing, the type and property of the seats, the sloping of tiers and the distribution of the persons in the room. In a diffuse sound field the location of the sound source towards the audience area is of minor importance in this regard. Fig. 7-50 shows a survey of the values of the equivalent sound absorption area per person for a variety of occupation densities and seating patterns in a diffuse sound field. Since in many types of rooms the reverberation time for medium and high frequencies is determined almost exclusively by the sound absorption of the audience, one has to reckon with a rather high error rate, if the range of dispersion of the factors influencing the sound absorption capacity of the audience is to be taken into account when determining the reverberation time, Fig. 7-50. A still wider range of dispersion of the sound absorption area occurs with the musicians and their instruments, Fig. 7-51. The unilateral arrangement of the listener or musician areas prevailing in most rooms tends to disturb the diffusivity of the sound field heavily so that the above-mentioned measured values may be faulty, Figs. 7-50 and 7-51.

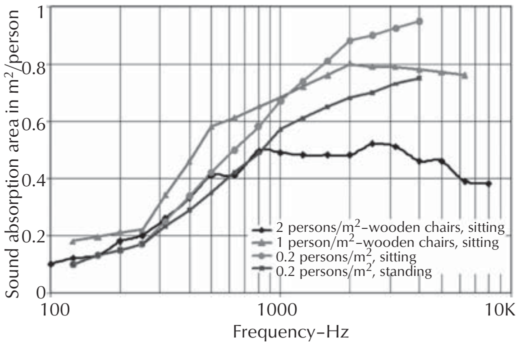


Figure 7-50. Equivalent sound absorption area in m^2/person of audience.

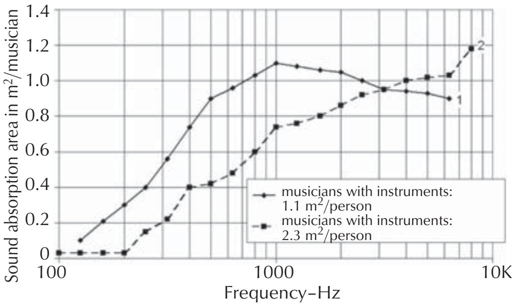


Figure 7-51. Equivalent sound absorption area in m^2/person of musicians.

Especially with an almost plain arrangement of the audience and performance areas there occurs for the direct sound and the initial reflections a frequency-dependent additional attenuation through the grazing sound incidence on the audience area. This is intensified by the fact that the sound receivers—i.e., the ears—are located in this indifferent acoustical boundary region, so that the influence of this additional attenuation becomes particularly relevant for the auditory impression. According to Mommertz[45] this effect of additional attenuation can be attributed to three causes:

1. The periodical structure of seat arrangement compels a guided wave propagation for low frequencies. In the frequency range between 150 Hz and 250 Hz, this additional attenuation causes a frequency-selective level dip which is designated as *seat dip effect*. An example is given in Fig. 7-52 for a frequency of about 200 Hz.[45]
2. The scattering of sound at the heads produces an additional attenuation especially in the frequency range between 1.5 kHz and 4 kHz, which is designated as *head dip effect*, Fig. 7-52. The magnitude of the effect depends largely on the seat arrangement and the orientation of the head with regard to the sound source.
3. In combination with the incident direct sound, the scattering at the shoulders produces a very broadband additional attenuation through interference. It is possible to define a simple correlation between the so-called elevation angle, Fig. 7-53, and the sound level reduction in the medium-frequency range at ear level of a sitting person.[45]

$$\Delta L = -20\log(0.2 + 0.1\gamma) \tag{7-72}$$

where,

ΔL is in dB,

γ is in degrees, $\gamma < 8°$.

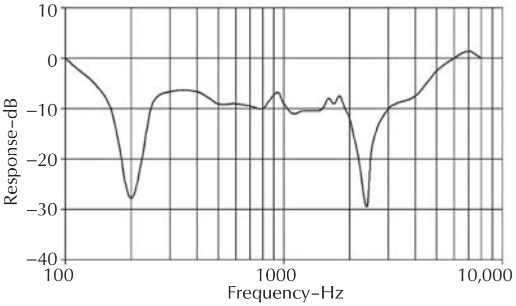
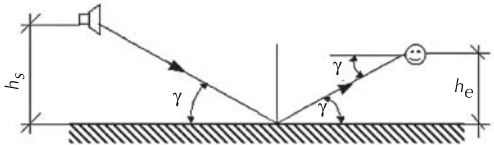


Figure 7-52. Measured quantity spectrum of the transfer function above a plain audience arrangement.[45]

Fig. 7-54 shows a graphical representation of the correlation resulting from Eq. 7-72. One sees that with a plain arrangement of source and receiver the resulting level reduction may be up to about 14 dB, whereas an elevation angle of 7° suffices to cut the level reduction to a negligible amount of less than 1 dB. The reflection plane defined in Fig. 7-53 lies herewith $h_L = 0.15$ m (0.5 ft) below ear level, for example, at shoulder level of a sitting person (approximately 1.05 m [3.5 ft] above the upper edge of floor). According to Reference 45, the additional attenuation depends herewith only on the elevation angle, no matter if tier sloping is effected in the audience area or in the performance area.



h_s = height of the sound source above the reflection plane
h_e = height of the receiver above the reflection plane
γ = elevation angle

Figure 7-53. Geometric data for determining the elevation angle above a sound reflecting plane.

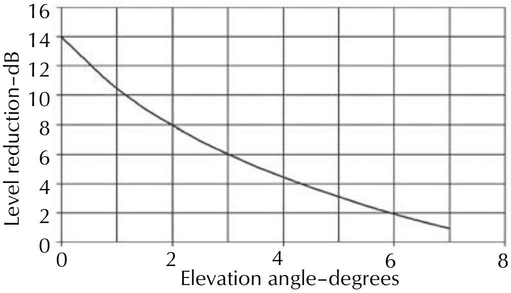


Figure 7-54. Sound pressure level reduction by sound dispersion at shoulder level of sitting persons as a function of the elevation angle.

Fig. 7-55 shows the influence of the height of the sound source above the ear level of a person sitting in a row at a distance source-receiver of 15 m (50 ft), on the frequency-dependent additional attenuation caused by a grazing sound incidence over an audience.[45] The receiver level is herewith 1.2 m (4 ft) above the upper edge of the floor, the height of the source above the floor is represented in this example as being 1.4 m and 2.0 m (4.6 ft and 6.6 ft). With a level difference of only 0.2 m (0.66 ft) between source and receiver, one can clearly recognize the additional timbre change of the direct sound component and of the initial reflections by attenuation in the low- and medium-frequency ranges, whereas with a level difference of 0.8 m (2.6 ft) the sound level attenuations get reduced to below 3 dB.

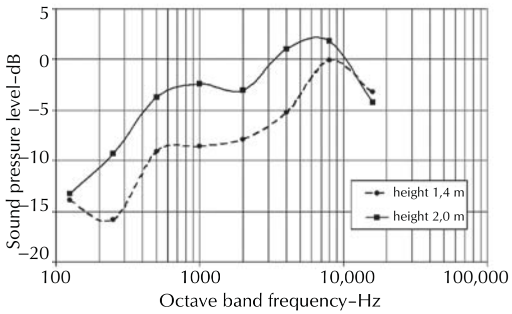


Figure 7-55. Influence of the vertical distance ear level-sound source level on the sound pressure level at a listener seat referred to free field radiation.

7.4 Variation of Room Acoustics by Construction or Electroacoustic Methods

7.4.1 Variable Acoustics

The manipulation of room acoustic properties is universally known by the term vario-acoustics as well as variable acoustics. What variable manipulations are possible in variable acoustics? The primary structures (volume, dimensions, shape) and the secondary structures (reflective, diffuse, absorptive) of a room have a huge influence on its acoustic properties.

Acoustical parameters describing the acoustical overall impression of rooms are determined by the utilization function (see Section 7.3.1). If this function is unambiguously defined, as is the case, for example, with auditoriums and concert halls for symphonic music, the result corresponds, provided an appropriate planning was carried out, to the utilization-relevant room-acoustical requirements. Things look quite different, however, for rooms having a wide utilization range—i.e., so-called multipurpose halls. For speech and music performances which use exclusively sound reinforcement, a short reverberation time with little rise in the low-frequency range as well as a reduced spaciousness of the natural sound field are desirable. For vocal and music performances with mainly natural sound sources, however, a longer reverberation time with enhanced spaciousness is aspired to. The timbre of the room should herewith show more warmth in the lower frequency range. As regards their room-acoustical planning, most multipurpose halls feature a compromise solution which is harmonized with their main utilization variant and does not allow any variability. The acoustically better solution lies in the realization of variable acoustics within certain limits. This aim can be achieved by architectural or electroacoustical means.

Another range of application for purposefully variable room-acoustical parameters is by influencing the reverberation time, clarity and spaciousness of rooms which owing to their form and size meet with narrow physical boundaries in this respect. This concerns mainly rooms with too small a volume index (see Section 7.3.2.1), or such containing a large portion of sound-absorbing materials. Architectural measures for achieving desirable modifications of room-acoustical parameters are applicable here only to a limited extent, since they are bound to implement methods allowing a deliberate modification of the temporal and frequency-dependent sound energy behavior of the sound field. The effectiveness of these methods is herewith determined by the correspondingly relevant sound energy component of the room-acoustical parameter. Achieving a desired reverberation-time and spaciousness enhancement requires a prolongation of the travel path of sound reflections and a reduction of the sound absorption of late sound reflections (enhancement of reverberant energy). In this respect, more favorable results can be obtained by electroacoustical means, particularly because in such rooms the sound-field structure does not contribute essentially to the manipulated parameters. From a practical point of view Section 7.4 is mainly dedicated to the presentation of electronic procedures for reverberation-time prolongation. Equivalent architectural measures will be explained only on fundamental lines.

7.4.1.1 Fundamentals of Variable Acoustics

In the planning and realization of measures enabling variation of room-acoustical parameters, it is necessary to comply with essential aspects so that the listener's subjective listening expectation in the room is not spoiled by an excessive range of variability:

1. The measures viable for realizing variable acoustics by architectural as well as electroacoustical means can be derived from the definitions of the room-acoustical parameters to be varied (see Section 7.2.2). Additional sound reflections arriving exclusively from the direction of the sound source surely enhance clarity, but boost spaciousness as little as an additional lateral sound energy prolongs reverberation time. Spaciousness-enhancing and reverberation-time prolonging sound reflections must essentially impact on the listener from all directions of the room. By means of appropriately dimensioned additional architec-

tural measures it is possible to achieve good results in this respect. Realization of the same, however, often implies high technical expenditure and costs. For influencing the reverberation time, for instance, these include the coupling or uncoupling of additional room volumes or the prolongation of the travel path of sound reflections with simultaneous reduction of the sound absorption of late sound reflections. A desired reduction of reverberation time and spatial impression can be achieved by means of variable sound-absorbing materials (curtains, slewable wall elements of different acoustic wall impedance) which have to be effective over the whole frequency range required by the performance concerned.

2. The coupling of acoustically effective room volumes has to be done in such a way that these acoustically form a unity with the original room. Otherwise there occur disturbing effects like timbre changes and double-slope reverberation-time curves. Incorrect dimensioning often results, owing to an acoustical orientation towards the additional room volume, in a heavily frequency-dependent spaciousness of the decay process in the sound field. The frequency-dependent reverberation time of the additional room volume must be a bit longer than or at least as long as that of the original room.

   In the opposite case of reducing the reverberation time by uncoupling the additional room volume, it is for the remaining room volume necessary to provide the sound-field structure required for the desired variation. For instance, there is more sound energy to be allocated to the initial reflections and in the decay process—which is now to be supplied with less sound energy—there must not occur any irregularities.

3. The variation depth achievable by means of variable acoustics must be acoustically perceptible to a significant degree. The distinctive threshold of, for example, subjectively perceived reverberation time changes is not independent of the absolute value of the reverberation time. Variations of 0.1 s to 0.2 s are at medium frequencies and a reverberation time of up to 1.4 s to 1.5 s is subjectively less clearly perceived than above this limit value. Thus a reverberation-time prolongation from 1.0 s to 1.2 s attained with much technical effort is almost not audible, whereas one from 1.6 s to 1.8 s is already significantly audible.
4. The listening experience has to tally with the overall visual impression of the room—too heavy deviations are perceived as disturbing and unnatural. This aspect has to be taken into account especially with small room volumes, if an excessively long reverberation time is produced by an electronic enhancement system (except for acoustic disassociation effects).
5. The sound-field structure of the original room has to remain unchanged if measures of variable acoustics are implemented. Additionally modified sound reflections have to correspond to the frequency and time structure of the room. This aspect holds true for architectural as well as electroacoustical measures—e.g., for reverberation enhancement. Coupled additional room volumes must not involve any distinctive timbre changes compared with the main room. Electroacoustical room sound simulators with synthetically produced sound fields are allowed to change the transmission function only in compliance with the original room, except if alienation effects are required for special play-ins.
6. An enhancement of reverberation time and spaciousness is possible only within permissible boundaries in which the overall acoustic impression is not noticeably disturbed. This boundary is all the lower the more the manipulation makes the sound field structure deviate from that of the original room.

Aspects to Be Considered with the Realization of Variable Acoustics. In keeping with the envisaged target, the following main realization objectives can be formulated for variable acoustics:

1. Large Room Volume (Large Volume Index) or Reverberant Rooms

   - **Task of variable acoustics:** Enhancement of clarity and definition. Reduction of reverberation time and spaciousness.
   - **Architectural solution:** Apart from an appropriate tiering arrangement of the sound sources, variable ceiling panels and movable wall elements have to be placed at defined distances for enhancing the clarity of music and the definition of speech. Modified inclinations of walls, built-in concert shells near stage areas in theaters, etc., create new primary reflections that are in harmony with the variants of purpose.

Broadband sound absorbers in the shape of variable mechanisms for low frequencies, combined with curtain elements or slewable boundary elements of differing acoustic wall impedance, reduce reverberation time and diminish spaciousness. When arranging variable sound absorbers it is necessary to pay attention to the frequency dependence of these elements. Slots between the installed slewable wall elements may, depending on the position of the elements, function as unwanted additional bass absorbers. In case of exclusive use of curtain arrangements, the low-frequency fraction is at a disadvantage giving rise to a brilliance-poor sound pattern.

An effective broadband reduction of the reverberation time can be achieved by deliberately influencing the late sound reflection mechanism. This may be realized by means of mobile room dividing wall parts which shorten the travel distances of sound reflections at the points of maximum length or width of the room and direct the reflections toward the sound-absorbing listener area. To this effect it is also possible to perform a room reduction, for example, by detaching the room volume above or below a balcony or a gallery.

- **Electronic solution:** An additional electronic architecture system serves for enhancing definition and clarity. Reducing reverberation time or diminishing spaciousness however, are not possible by electronic means.

2. Little Room Volume (Small Volume Index) or Highly Sound-Absorbent Rooms

   - **Task of variable acoustics:** Enhancement of reverberation time and spaciousness.
   - **Architectural solution:** One solution for enhancing reverberation time consists of the coupling of additional room volumes in acoustical unity with the main room. By means of a purposive sound reflection guidance at the point of maximum room length or width, it is possible to realize long travel paths by letting the sound repeatedly be reflected between the room boundary faces and the room ceiling, thus having it belatedly absorbed by the listener area (cf. large concert hall in the Neues Gewandhaus Leipzig). This way it is first of all the early decay time, which is mainly responsible for the subjectively perceived reverberation duration, which is prolonged.
   - **Electronic solution:** Influencing reverberation time and spaciousness is quite possible by electronic means, if the physical and auditory-psychological limitations are observed. Viable solutions are described in detail in Section 7.4.2.

In general variable-acoustics is steadily losing ground because of its high costs and low effect in comparison with the use of correctly designed sound systems.

7.4.2 Electronic Architecture

Establishing good audibility in rooms, indoors as well as in the open air, has been and remains the object of room acoustics. This is the reason why room acoustics is called architectural acoustics in some countries.

The architectural measures have far-reaching limitations. These shortcomings are:

- The sound source in question has only a limited sound power rate.
- Changes of room acoustics may make huge changes in the architectural design and thus cannot always be optimally applied.
- The measures regarding room acoustics may make a considerable amount of constructional changes and these can only be done optimally for one intended purpose of the room.
- The constructional change, despite its high costs, results in only a very limited effect.

Because of these reasons sound systems are increasingly being used to influence specific room acoustic properties, thus improving audibility. This holds true regarding an improvement of intelligibility as well as of spaciousness. So one can speak of a good acoustic design if one cannot distinguish, when listening to an event, whether the sound quality is caused only by the original source or by using an electroacoustic installation.

Another task of sound installation techniques consists of electronically modified signals that remain uninfluenced by specific properties of the listener room. It is necessary to suppress, as far as possible, the acoustic properties of the respective listener room by using directed loudspeaker systems. It is also possible to create a dry acoustic atmosphere by using suitable supplementary room-acoustic measures.

Reverberation (the reverberation time of a room) cannot be reduced by means of sound systems. At the typical listener seat the level of the direct sound is of great significance. Also short-time reflections, enhancing intelligibility of speech and clarity of music, can be provided by means of sound reinforcement.

The following sound-field components can be manipulated or generated:

- Direct sound.
- Initial reflections with direct-sound effect.
- Reverberant initial reflections.
- Reverberation.

For this reason electronic techniques were developed that opened up the possibility of increasing the direct or reverberation time and energy in the hall—i.e., directly influencing the acoustic room properties.

Such a method for changing the room-acoustic properties of rooms is now called the application of electronic architecture.

7.4.2.1 Use of Sound Delay Systems for Enhancing Spaciousness

These procedures act in particular on the sound energy of the initial reflections affecting the reverberant sound.

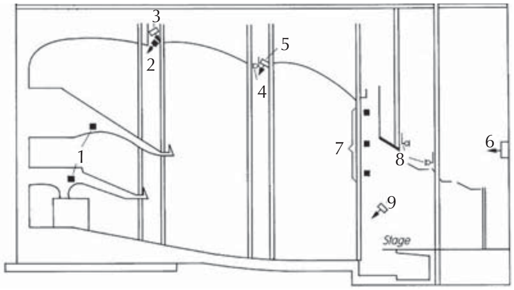
7.4.2.1.1 Ambiophony

This procedure, which is already obsolete, makes use of delaying devices reproducing not only the discrete initial reflections; but also the reverberation tail. The reflection sequences have herewith to be chosen in such a way that no comb-filter effects, such as flutter echoes, will be produced with impulsive music motifs. The functioning of a simple ambiophonic system can be described as follows: to the direct sound emanating directly from the original source and directly irradiated into the room, there are admixed delayed signals produced by an adequate sound delaying system (in the initial stages this was just a magnetic sound recording system) which are then irradiated like reflections arriving with corresponding delay from the walls or the ceiling. This requires additional loudspeakers appropriately distributed in the room for irradiating the delayed sound as diffusely as possible. For further delaying the sound it is possible to arrange an additional feedback from the last output of the delay chain to the input. A system of this kind was first suggested by Kleis[51] and was installed realized in several large halls.[52,53]

7.4.2.1.2 ERES (Electronic Reflected Energy System)

This procedure was suggested by Jaffe and is based on a simulation of early reflections used for producing so-called reverberant-sound-efficient initial reflections, Fig. 7-56.[54]

Thanks to the arrangement of the loudspeakers in the walls of the stage-near hall area and to the variability range available for delay, filtering and level regulation of the signals supplied to them, adapted lateral reflections can be irradiated. The spatial impression can thus be amply influenced by simulating an acoustically wider portal by means of a longer reverberation time or a narrower portal by using a shorter reverberation time.



1. One of the 14 pairs of the AR (assisted resonance)/ ERES loudspeakers under the balcony.
2. One of the 90 AR loudspeakers.
3. One of the four ERES loudspeakers in the third ceiling offset.
4. One of the 90 AR microphones.
5. One of the four ERES loudspeakers in the second ceiling offset.
6. ERES stage-tower loudspeaker.
7. Three of the six AR proscenium loudspeakers.
8. ERES microphones.
9. One of the two ERES proscenium loudspeakers.

Figure 7-56. ERES/AR system in the Sivia Hall in the Eugene Performing Arts Center, Eugene, Oregon.

This gives the capability of:

- Adaptation to acoustical requirements.
- Simulation of different hall sizes.
- Optimization of definition and clarity.

Jaffe and collaborators speak of electronic architecture. It is certainly true that this selective play-in of reflections does simulate room-acoustical properties the room in question is devoid of, so as to compensate shortcomings in its room-acoustical structure. After installing the first system of this kind in the Eugene Performing Arts Center in Oregon,[55] Jaffe-Acoustics have installed further ones in a large number of halls in the United Statres, Canada and other countries.

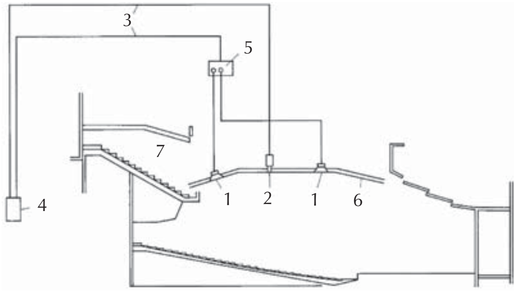
The electronic delay procedure in sound reinforcement systems has meanwhile become a general practice all over the world and is now the standard technique used for the play-in of delayed signals (e.g., for simulating late reflections). In this sense one may well say that electronic architecture is used in all instances where such reflections are used on purpose or unintentionally.

7.4.2.2 Travel-Time-Based Reverberation-Time Enhancing Systems

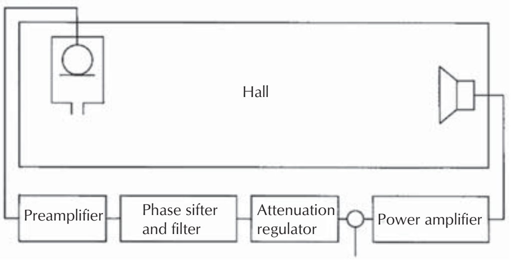
This procedure is mainly used for enhancing the late reverberant sound energy combined with an enhancement of the reverberation time.

7.4.2.2.1 Assisted Resonance

For optimizing the reverberation time in the Royal Festival Hall built in London in 1951, Parkin and Morgan[56,57] suggested a procedure permitting an enhancement of the reverberation time especially for low frequencies, Fig. 7-57.



A. Principal layout of a channel arrangement in a hall.



B. Components of an AR-Channel (Microphone in resonance chamber).

1. 60 loudspeaker boxes each in the ceiling area and in the upper wall region.
2. 120 microphones in Helmholtz resonator boxes.
3. 120 microphone and loudspeaker cables.
4. Remote control, phase shifter, amplifier for the 120 channels.
5. Distributor for loudspeaker boxes.
6. Movable ceiling for varying the volume of the room.
7. Balcony.

Figure 7-57. Assisted resonance system.

Parkin and Morgan proceeded on the assumption that in any room there exist a multitude of eigenfrequencies which give rise to standing waves with nodes and antinodes decaying by an e-function according to the absorption characteristics of the surface involved. This decay process is characteristic for the reverberation time of the room at the corresponding frequency. Any standing wave has its specific orientation in space and a microphone is installed at the point where a sound pressure maximum (vibration antinode) occurs for a given frequency. The energy picked up from the microphone is supplied via an amplifier to a loudspeaker installed at a distant antinode of the same standing wave, so that the energy lost by absorption is compensated. The energy at that frequency can thus be sustained for a longer period (assisted resonance). By enhancing the amplification it is possible to considerably prolong the reverberation time for this frequency (until feedback sets in). Thanks to the spatial distribution of the irradiating loudspeakers this applies accordingly to the spatial impression.

These considerations hold true for all eigenfrequencies of the room. The arrangement of the microphones and loudspeakers at the locations determined by the antinodes of the individual eigenfrequencies may, however, be difficult. The microphones and loudspeakers are therefore installed at less critical points and driven via phase shifters. In the transmission path there are additionally inserted filters (Helmholtz resonators, bandwidth approximately 3 Hz) which allow the transmission channel to respond only at the corresponding eigenfrequency. Care should be taken that the irradiating loudspeakers are not arranged at a greater distance from the performance area than their corresponding microphones, since the first arrival of the reverberant signal may produce mislocalization of the source.

This procedure, which has meanwhile become obsolete, was installed in a large number of halls. In spite of its high technical expenditure and the fact that the system required can be used only for the assisted resonance, it was for a long period one of the most reliable solutions for enhancing the reverberation time without affecting the sound, particularly at low frequencies.

7.4.2.2.2 Multi-Channel-Reverberation, MCR

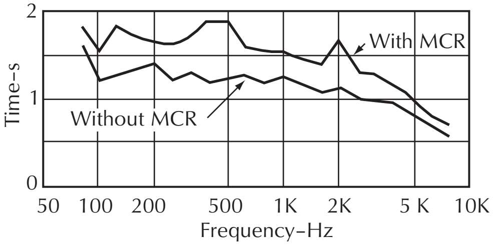
Using a large number of broadband transmission channels whose amplification per channel is so low that no timbre change due to commencing feedback can occur, was suggested first by Franssen.[58] While the individual channel remaining below the positive feedback threshold provides only little amplification, the multitude of channels is able to produce an energy density capable of notably enhancing the spatial impression and the reverberation time.

The enhancement of the reverberation time is determined by

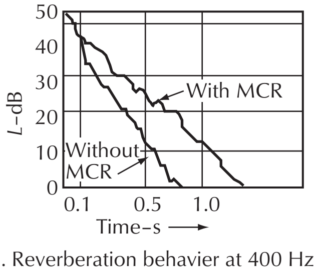
$$\frac{T_m}{T_o} = 1 + \frac{n}{50} \tag{7-73}$$

If the reverberation time is, for instance, to be doubled (which means doubling the energy density), there are $n = 50$ individual amplification chains required. Ohsmann[59] has investigated in an extensive paper the functional principle of these loudspeaker systems and has shown that the prognosticated results regarding enhancement of the reverberation time cannot be achieved in practice. He also quotes the fact that Franssen "did not sufficiently consider the cross couplings between the channels" as a possible reason for the deviation from theory.[1]

A system technologically based on this procedure is offered by Philips under the name of Multi-Channel Amplification of Reverberation System (MCR). It serves for enhancing reverberation and spaciousness.[60] According to manufacturer's specifications a prolongation of the average reverberation time from approximately 1.2 to 1.7 s is achieved for ninty channels. Even longer reverberation enhancements are said to be possible. There exist numerous implementations in medium-sized and large halls (the first was in the POC Theater in Eindhoven, Fig. 7-58).



A. Frequency response of the reverberation time with and without MCR.



B. Reverberation behavier at 400 Hz.

Technical data of the system: hall 3100 m^3, stage 900 m^3.
90 channels (preamplifier, filter, power amplifier).
90 microphones at the ceiling.
110 loudspeakers in the side walls, in the ceiling and under the balcony.
Remote control of the reverberation in 10 steps.

Figure 7-58. MCR system in the POC Theater in Eindhoven.

7.4.2.3 Modern Procedures for Enhancing Reverberation and Spaciousness

7.4.2.3.1 Acoustic Control System (ACS)

This procedure was developed by Berkhout and de Vries at the University of Delft.[61] Based on a wave-field synthesis approach (WFS) the authors speak of a holographic attempt for enhancing the reverberation in rooms. In essence, it is really more than the result of a mathematical-physical convolution of signals captured by means of microphones in an in-line arrangement (as is the case with WFS). The room characteristics are predetermined by a processor, which produces, in the end, a new room characteristic with a new reverberation time behavior, Fig 7-59.

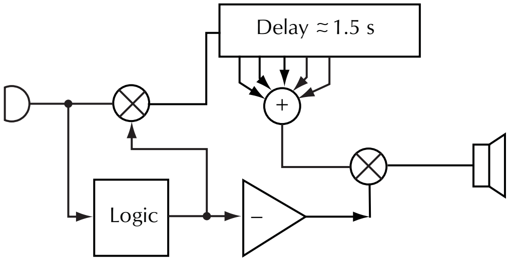


Figure 7-59. Principle of the Reverberation on Demand System (RODS).

The upper block diagram shows the principle of the ACS circuit for a loudspeaker-microphone pair. One sees that the acoustician formulates the characteristics of a desired room—e.g., in a computer model—transfers these characteristics by means of suitable parameters to a reflection simulator and convolutes these reflection patterns with the real acoustical characteristics of a hall. Fig. 7-60 shows the complete block diagram of an ACS system.

Unlike other systems, the ACS does not use any feedback loops—thus timbre changes owing to self-excitation phenomena should not be expected. The system is functioning in a series of halls in the Netherlands, Great Britain and in the United States.

7.4.2.3.2 Reverberation on Demand System, RODS

With this system a microphone signal is picked near the source and passed through a logical switching gate before reaching a delay line with branched members. This output is equipped with a similar gate. A logical control circuit opens the input gate and closes the output gate when the microphone signal is constant or rising.

Vice versa, it closes the input gate and opens the output gate when the microphone signal is falling, Fig. 7-60.[62]

An acoustical feedback is thus avoided, but this system fails to enhance the lateral energy with continuous music, which makes it unsuitable for music performances. It is no longer used.

7.4.2.3.3 LARES

The LARES system by the Lexicon Company uses modules of the standardized Room Processor 480L which, when fed by special software, simulates the desired decay curves, Fig. 7-61. A large number of loudspeakers are required in the wall and ceiling areas. The input signals are picked up by just a few microphones in an area near the source.[63,64] On account of the time-variant signal processing (a large quantity of independent time-variant reverberation devices), the adjustment of reverberation times is not exactly repeatable. Common computer-controlled measuring software (based, e.g., on MLS) is thus unable to measure decay curves. Apart from the ASC system, LARES installations are very widespread in Europe and the United States. Well known are the systems installed in the Staatsoper Berlin, the Staatsschauspiel Dresden, and the Seebühne (floating stage) in Mörbisch/Austria.

7.4.2.3.4 System for Improved Acoustic Performance (SIAP)

The basic principle of SIAP consists in picking up the sound produced by the source by means of a relatively small number of microphones, processing it appropriately (by means of processors which convolute, that is overlay electronically the room-acoustical parameters of a room with target parameters) and then feeding it back into the hall by an adequate number of loudspeakers, Fig. 7-62. The aim is to produce desired natural acoustical properties by electronic means. For obtaining spatial diffusivity a large number of different output channels are required. Moreover, the maximally attainable acoustic amplification is dependent on the number of uncorrelated paths. Compared with a simple feedback channel, a system with 4 inputs and 25 outputs is able to produce a 20 dB higher amplification before feedback sets in. This holds true, of course, only under the assumption that each and every input and output path is sufficiently decoupled from the other input/output paths. Each listener seat receives sound from several loudspeakers, each of which irradiates a signal somewhat differently processed than any of the others (!).[65]

7.4.2.3.5 Active Field Control, AFC

The AFC system by Yamaha[66] makes active use of acoustic feedback for enhancing the sound energy density and thereby also the reverberation time. When using the acoustic feedback it is, however, important to avoid timbre changes and to insure the stability of the system. To this effect one uses a specific switching circuit, the so-called Time Varying Control (TVC) which consists of two components:

- Electronic Microphone Rotator (EMR).
- Fluctuating FIR (fluc-FIR).

The EMR unit scans the boundary microphones in cycles while the FIR filters impede feedback.

For enhancing the reverberation, the microphones are arranged in the diffuse sound field and still in the close-range source area (gray dots in Fig. 7-63 on the right). The loudspeakers are located in the wall and ceiling areas of the room. For enhancing the early reflections there are four to eight microphones arranged in the ceiling area near the sources. The signals picked up by these are passed through FIR filters and reproduced as

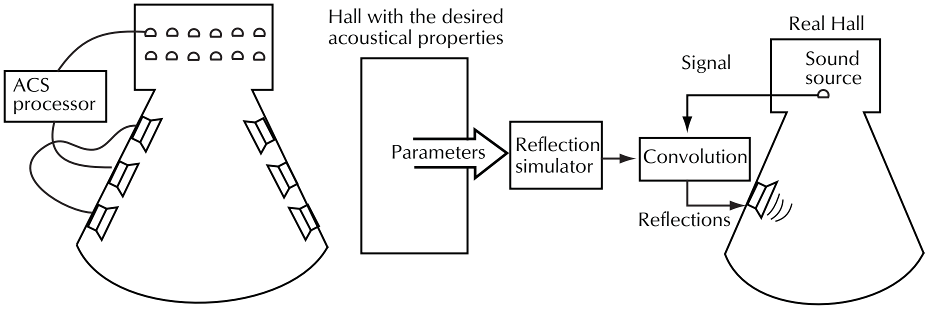


Figure 7-60. Basic block diagram of the Acoustic Control System (ACS) illustrated for a loudspeaker–microphone pair.

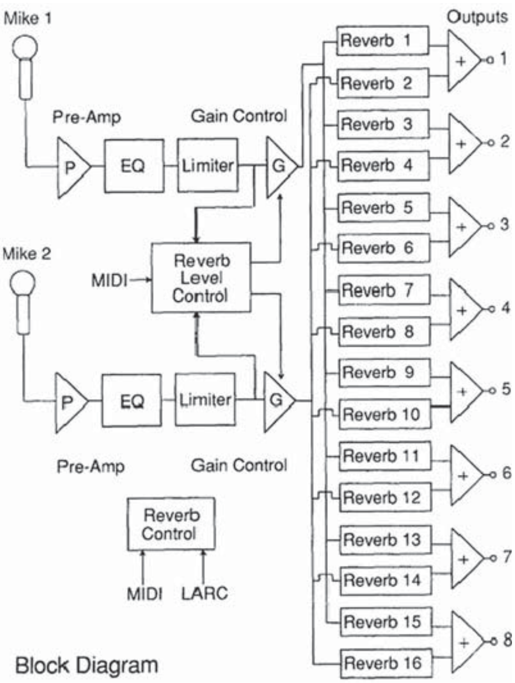

Mike 1
Outputs
Pre-Amp
Gain Control
P
EQ
Limiter
G
Mike 2
MIDI
Reverb Level Control
P
EQ
Limiter
G
Pre-Amp
Gain Control
Reverb Control
MIDI
LARC
Reverb 1
Reverb 2
Reverb 3
Reverb 4
Reverb 5
Reverb 6
Reverb 7
Reverb 8
Reverb 9
Reverb 10
Reverb 11
Reverb 12
Reverb 13
Reverb 14
Reverb 15
Reverb 16
1
2
3
4
5
6
7
8
Block Diagram

Figure 7-61. LARES block diagram.

lateral reflections by loudspeakers located in the wall and ceiling areas of the room. The loudspeakers are arranged in such a way that they cannot be located, since their signals are to be perceived as natural reflections.

Furthermore the AFC system allows signals to be picked up in the central region of the audience area and the reproduction of them via ceiling loudspeakers in the area below the balcony for the sake of enhancing spaciousness.

7.4.2.3.6 Virtual Room Acoustic System Constellation™

Constellation™ by Meyer Sound Inc. is a multichannel regenerative system for reverberation enhancement. Its development is based on ideas already considered in the sixties of last century by Franssen[58] when developing the above-mentioned MCR procedure.[67] The principle is already rather old and was described by users.[68] The biggest difference is that today Constellation™ uses, instead of the second reverberation room, an electronic reverberation processor which is more easily adaptable.

But modern electronic elements and DSPs have made it possible to design circuits which widely exclude timbre changes. This is achieved by coupling a primary room A (the theater or concert hall) with a secondary room B (the reverberant room processor). Simultaneously the number of reproduction channels is reduced along with the timbre change of sound events. An enhancement of the early reflections is obtained as well, cf. Fig. 7-64.

Contrary to other systems, Constellation™ uses a comparable number of microphones and loudspeakers in a room. To this effect the microphones are located in the reverberant or diffuse field of all sound sources within the room and connected via preamplifiers to a digital processor. Then the outputs of the processor are connected to power amplifiers and loudspeakers for reproduction of the signal.

With the Constellation™ system there is a multitude of small loudspeakers L_1 to L_N (40 to 50) distributed in

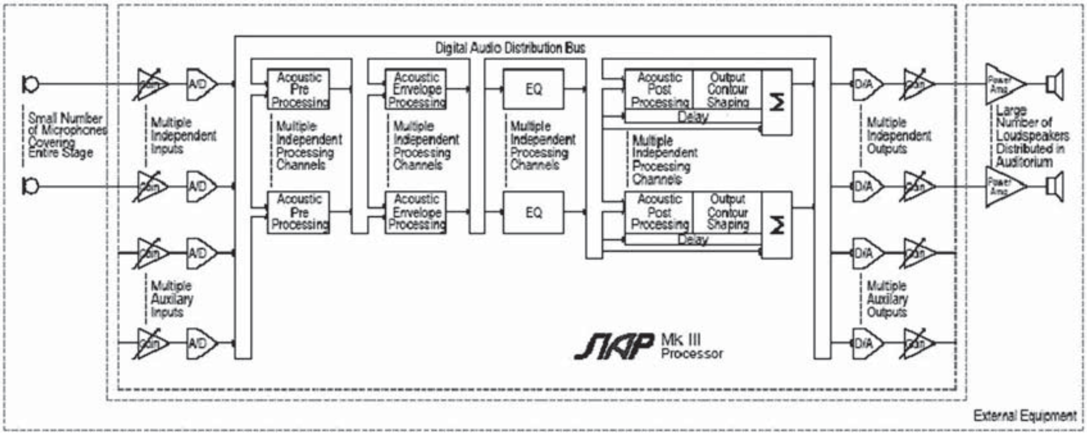

Digital Audio Distribution Bus
Small Number of Microphones Covering Entire Stage
Gain
A/D
Multiple Independent Inputs
Multiple Auxiliary Inputs
Acoustic Pre Processing
Multiple Independent Processing Channels
Acoustic Envelope Processing
EQ
Acoustic Post Processing
Output Contour Shaping
Delay
Σ
D/A
Multiple Independent Outputs
Multiple Auxiliary Outputs
Power Amp
Large Number of Loudspeakers Distributed in Auditorium
SIAP Mk III Processor
External Equipment

Figure 7-62. Schematic circuit diagram of SIAP.

the room, which, of course, may also be used for panorama and effect purposes. Ten to fifteen strategically located and visually inconspicuous microphones m_1 to m_N pick up the sound and transmit it to the effect processor $X(\omega)$ where the desired and adjustable reverberation takes place. The output signals thus obtained are fed back into the room. The advantage of this solution lies in the precise tuning of the reverberation processor enabling well-reproducible and thus also measurable results.

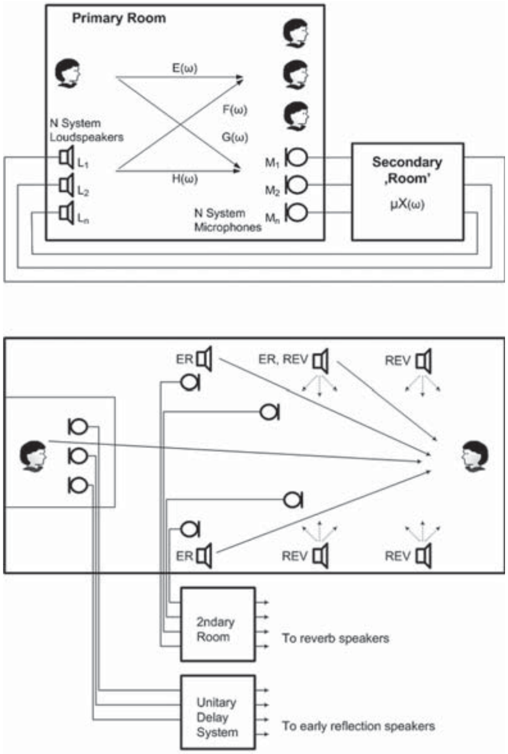

Primary Room
E(ω)
F(ω)
G(ω)
H(ω)
N System Loudspeakers
L₁
L₂
Lₙ
M₁
M₂
Mₙ
N System Microphones
Secondary 'Room'
µX(ω)
ER
ER, REV
REV
REV
ER
REV
REV
2ndary Room
To reverb speakers
Unitary Delay System
To early reflection speakers

Figure 7-64. Principle of the Virtual Room Acoustic System, Constellation™.

7.4.2.3.7 CARMEN®

The underlying principle is that of an active wall whose reflection properties can be electronically modified.[69] The system was called CARMEN® which is the French abbreviation of Active Reverberation Regulation through the Natural Effect of Virtual Walls. On the wall there are arranged so-called active cells forming a new virtual wall. The cells consist of a microphone, an electronic filter device, and a loudspeaker by which the picked-up signal is irradiated, Fig. 7-65. The microphones are typically located at 1 m distance from the loudspeaker of the respective cell, i.e. at approximately ⅕ of the diffuse-field distance in typical halls. Therefore it might also be called a locally active system.

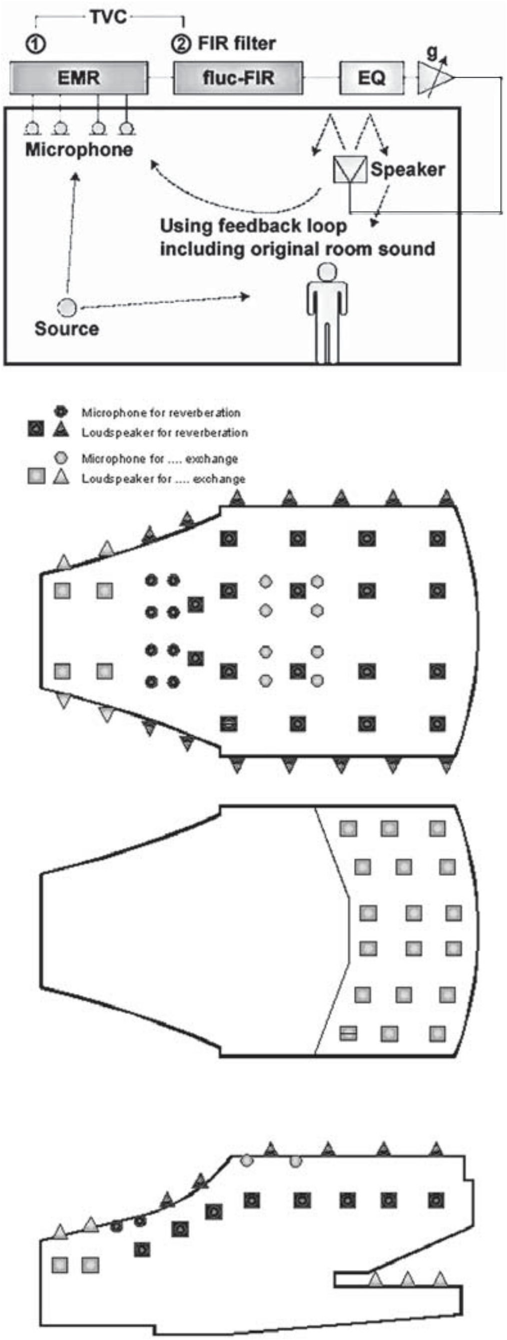

TVC
①
② FIR filter
EMR
fluc-FIR
EQ
g
Microphone
Speaker
Using feedback loop including original room sound
Source
Microphone for reverberation
Loudspeaker for reverberation
Microphone for .... exchange
Loudspeaker for .... exchange

Figure 7-63. Active Field Control System (AFC) by Yamaha, Japan.

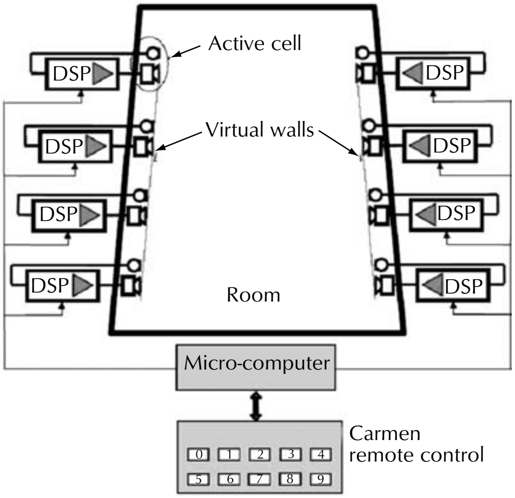


Figure 7-65. Principle of the Active Reverberation Regulation System, CARMEN®.

Every cell produces a desired decay of the artificial reflections, provided one does not choose an excessive cell gain liable to provoke feedback. To avoid feedback an adequate microphone directivity characteristic as well as internal echo canceling algorithms are used. In addition it is possible to delay the microphone signal electronically, a feature which allows the cell to be virtually shifted and the room volume to be apparently enlarged.

Since 1998 CARMEN® has been installed and tested in more than ten halls used by important orchestras. It has proven to be particularly effective in theaters which are also used for orchestra performances. In these rooms it best improves the acoustics in the distant areas under the balconies. In the Mogador Theater in Paris acoustics were significantly improved by installing CARMEN® cells in the side walls and in the ceiling of the balcony.

By means of twenty four cells in a room with a reverberation time of 1.2 s at 500 Hz, it was possible to enhance this reverberation time to 2.1 s. Additionally there resulted various spatial effects like a broadening of the sound source or an improved envelopment with lateral reflections, features often required for big orchestras, but also soloists.

7.4.3 Conclusions and Outlook

The above presented comparison shows that a large number of procedures exist for enhancing reverberation and spaciousness, part of which continue to be used today. Thanks to ever-increasing quality of the electronic transmission devices, the prejudices still existing especially among musicians against the electronic architecture will diminish, so that it will be increasingly possible to adapt concert halls even to the acoustical conditions characteristic of different creative and historic periods. Utilization in so-called multipurpose halls will, of course, prevail. The aim will have been achieved when musicians and audience perceive acoustical conditions established by means of electronic architecture as normal and natural. Simplicity of varying settings or security against acoustical feedback and unrelated timbre change will then be decisive factors in the choice of an enhancement system. Modern computer simulation will assist in banning the potential feedback risk.

Costly architectural measures for realizing the variable acoustic will be more and more discarded, particularly in view of their limited effectiveness.

References

1. W. Ahnert and W. Reichardt, *Grundlagen der Beschallungstechnik (Fundamentals of Sound Reinforcement Engineering),* S. Hirzel Verlag Stuttgart, 1981.
2. *Taschenbuch der Akustik (Handbook of Acoustics)*, Editors: W. Fasold, W. Kraak, and W. Schirmer. Verlag Technik, Berlin, 1984.
3. J. Meyer, *Akustik und musikalische Aufführungspraxis (Acoustics and Music Presentation Practice)*, Verlag Erwin Bochinsky, Frankfurt Am Main, 1995.
4. W.C. Sabine, *Collected Papers on Acoustics*, Cambridge: Harvard Univ. Press, 1923.
5. H. Kuttruff, *Room Acoustics*, 3rd Edition, Applied Science Publishers Ltd ,London, 1991.
6. W. Ahnert and W. Steffen, *Sound Reinforcement Engineering, Fundamentals and Practice*, E&fn Spon, London, 2000.
7. L.L. Beranek, *Music, Acoustics and Architecture*, New York: Wiley, 1962.
8. W. Fasold and E. Veres, *Schallschutz + Raumakustik in der Praxis (Noise Control and Room Acoustics in Practice)*, Verlag fuer Bauwesen Berlin, 1998.

9. L.L. Beranek, *Concert and Opera Halls: How They Sound,* Woodbury: Acoustical Society of America, 1996.
10. ISO 3382, Messung der Nachhallzeit von Räumen mit Hinweisen auf andere akustische Parameter (Measuring the Reverberation Time of Rooms, with Hints for Other Parameters) (2000–03).
11. Y. Ando, *Architectural Acoustics*, Springer-Verlag edition New York, Inc., 1998.
12. J. Hoffmeier, Untersuchungen zum Einfluß von Raumklangfärbungen auf die Deutlichkeit von Sprache (Investigations on the Impact of Surround Sound Coloration on Speech Definition), Diploma thesis, Technical University, Diplomarbeit An Der Tu, Dresden, 1996.
13. L. Dietsch, "Objektive raumakustische Kriterien zur Erfassung von Echostörungen und Lautstärken bei Sprach und Musikdarbietungen" (Objective Room-Acoustical Criteria for Capturing Echo Disturbances and Loudness Variations with Speech and Musical Performing), Dissertation Technische Universität Dresden, 1983.
14. R. Hoehne and G. Schroth, "Zur Wahrnehmbarkeit von Deutlichkeits- und Durchsichtigkeitsunterschieden in Zuhörersälen" (Perceptibility of Definition and Clarity Differences in Audience Halls). *Acustica 81* (1995), pp. 309–319.
15. T. Houtgast and T. Steeneken, H.J.M., "A review of the MTF Concept in Room Acoustics and its use for Estimating Speech Intelligibility in Auditoria" *J. Acoust. Soc. Amer.* 77 (1985), pp. 1060–1077.
16. "Rasti-Sprachübertragungsmesser Typ 3361" (Speech Transmission Meter Type 3361), Datenblatt Fa. Brüel & Kjær.
17. V. M. A. Peutz, "Articulation Loss of Consonants as a Criterion for Speech Transmission in a Room," *J. Audio Eng. Soc. 19* (1971), H 11, pp. 915–919.
18. W. Klein, "Articulation Loss of Consonants as a Basis for the Design and Judgement of Sound Reinforcement Systems," *J. Audio Eng. Soc. 19* (1971).
19. Technical Report ISO Tr 4870, Acoustics: The Construction and Calibration of Speech Intelligibility Tests, 1991.
20. V.O. Knudsen, *Architectural Acoustics*. New York: John Wiley and Sons, Inc., 1932.
21. W. Reichardt, O. Abdel Alim and W. Schmidt, "Definitionen und Meßgrundlage eines objektiven Masses zur Ermittlung der Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei Musikdarbie- tungen" (Definitions and Measuring Base of an Objective Measure for Ascertaining the Boundary Between Useful and Useless Transients at Musical Performances, *Acustica* 32 (1975) 3, p. 126.
22. W. Reichardt, O. Abdel Alim and W. Schmidt, "Zusammenhang zwischen Klarheitsmaß C und anderen objektiven raumakustischen Kriterien" (Correlation Between the Clarity Measure C and Other Objective Room-Acoustical Criteria). (1975), H 2, pp. 144-155.
23. W. Schmidt, "Raumakustische Gütekriterien und ihre objektive Bestimmung durch analoge oder digitale Auswertung von Impulsschalltestmessungen,"Dissertation B Technische Universität Dresden, 1979.
24. W. Kuhl, "Räumlichkeit als Komponente des Raumeindrucks" (Spaciousness as a Component of Spatial Impression). *Acustica* 40 (1978), p. 167 ff.
25. U. Lehmann, "Untersuchungen zur Bestimmung des Raumeindrucks bei Musikdarbietungen und Grundlagen der Optimierung" (Investigations on the Determination of Spatial Impression with Music Presentations and Bases for Optimization, Technische Universität Dresden, 1974.
26. W. Schmidt, "Zusammenhang zwischen Hallabstand und Nachhallzeit für den Raumeindruck (Correlation Between Direct to Reverberent Ratio and Reverberation Time with Regard to Spatial Impression). *Zeitschrift fuer Hochfrequenztechnik Und Elektroakustik* 77 1968, p. 37 ff.
27. M. Kleiner, "A New Way of Measuring Lateral Energy Fractions." *App. Acoust.*, Vol. 27, p. 321, (1989).
28. W. Reichardt and A. Kussev, "Ein-und Ausschwingvorgang von Musikinstrumenten und integrierte Hüllkurven ganzer Musikinstrumentengruppen Eines Orchesters," *Zeitschrift Für Elektrische Informations-und Energietechnik* 3 (1973) 2, pp. 73–88.
29. W. Fasold, E. Sonntag and H. Winkler, *Bauphysikalische Entwurfslehre: Bau-und Raumakustik (Theory of Structural-Physical Design, Architectural and Room Acoustics*), Verlag für Bauwesen Berlin, 1987.
30. J. Meyer, "Zur Dynamik und Schalleistung von Orchesterinstrumenten," (On the Dynamic Range and Acoustic Power of Orchestra Instruments). *Acustica* 71 (1990), pp. 277–278.
31. H. Winkler and U. Stephenson, "Einfluss geneigter Seitenwände auf die Nachhallzeit" (The Influence of Inclined Side Walls on Reverberation Time), Fortschritte der Akustik - DAGA, (1990), pp. 867–874.
32. L. Cremer, "Early Lateral Reflections in Some Modern Concert Halls," *J. Acoust. Soc. Am. 85,* (March 1989) 3.

33. H.P. Tennhardt, “Modellmessverfahren für Balanceuntersuchungen bei Musikdarbietungen am Beispiel der Projektierung des grossen Saales im Neuen Gewandhaus Leipzig” (Model Measuring Procedures for Balance Investigations of Musical Performances Demonstrated on the Example of the great Hall of the Neues Gewandhaus Leipzig), *Acustica* 56 (1984) Nr. 2 (Oktober 1984), pp. 126–135.
34. L. Cremer and H. Mueller, *Die Wissenschaftlichen Grundlagen der Raumakustik, Band I (The Scientific Foundations of Room Acoustics*, Vol. 1, S. Hirzel Verlag Stuttgart, 1978.
35. H. Winkler, “Das Nachhallreservoir, Bedeutung und Beeinflussung” (The Reverberation Time Reservoir, Importance and Control), *Fortschritte der Akustik—DAGA* 95, pp. 315-318.
36. K. Takaku and S. Nakamura, “A Report on the Relationship between Orchestra Shell Design and Musicians’ Acoustical Impression,” 1*5th Intern. Congress on Acoustics*, Trondheim, June 1995, pp. 525–528.
37. A. Askenfeld, Stage Floors and Risers—Supporting Resonant Bodies or Sound Traps? In: S. Ternström, *Acoustics for Choirs and Orchestra.* Stockholm, 1986.
38. J.H. Rindel, Acoustic Design of Reflectors in Auditoria. Proc. Inst. of Ac. 14, Part 2, p. 119.
39. M. Vorlaender, and E. Mommertz, “Definition and Measurement of Random-Incidence Scattering Coefficients,” *Appl. Acoustics 60*, (2000), (2), 187-200 .
40. E. Meyer and L. Bohn, “Schallreflexion an Flächen mit periodischer Struktur” (Sound Reflections at Surfaces with Periodical Structure), *Acustica* 2 (1952), H. 4, pp. 195–207.
41. M. R. Schroeder, “Binaural Dissimilarity and Optimum Ceilings for Concert Halls: More Lateral Sound Diffusion,” *JASA*, Vol. 65 (1992), No.4, p. 958 ff.
42. M. R. Schroeder, “Modulation Transfer Functions: Definition and Measurement,” *Acustica*, Vol. 49 (1981).
43. N. Kiesewetter, “Schallabsorption durch Platten-Resonatoren” (Sound Absorption by Panel Resonators), *Gesundheits-Ingenieur Gi 101*, (1980), H.3, pp. 57–62.
44. The New Microsorber from Käfer, Product Information of Käfer Isoliertechnik, Bremen, Germany, www.kaefer.com.
45. E. Mommertz, *Untersuchungen akustischer Wandeigenschaften und Modellierung der Schallrückwürfe in der binauralen Raumsimulation, (Investigations of Acoustic Wall Properties and Modelling of Sound Reflections in Binaural Room Simulations)*, Aachen: Shaker, 1996.
46. L.L. Beranek, *Concert halls and Opera Houses, Music, Acoustics, and Architecture*, New York: Springer, 2nd edition 2004.
47. T. Cox and P. D’Antonio, *Acoustic Diffusers and Absorbers: Theory, Design and Application*, London: T&F Group, SPON, 2004.
48. M. Vorlaender, and E. Mommertz, “Measurement of Scattering Coefficients of Surfaces in the Reverberation Chamber and in the Free Field,” *Proc. 15th ICA*, II, (1995), 577–580.
49. M. H. Gomes, M. Vorlaender, and S.N. Gerges, “Anforderungen an die Probeflächengeometrie bei der Messung des Streugrades im Diffusfeld,” *Fortschritte der Akustik—DAGA* ’02 – Bochum, S. 584–585.
50. ISO 17497, 2004 (E), Acoustics—Sound-Scattering Properties of Surfaces.
51. D. Kleis, “Moderne Beschallungstechnik” (Modern Sound Reinforcement). Philips tech. Rdsch. 20 (1958/59) 9, S. 272ff. und 21 (1959/60) 3, S. 78ff.
52. E. Meyer, and H. Kuttruff, “Zur Raumakustik einer großen Festhalle” (Regarding the Room Acoustic of a Large Festival Hall), *Acustica 14* (1964) 3, S. 138–147.
53. W. Kaczorowski, “Urzadzenia elektroakustyczne i naglosnia w Wojewodzkiej Hali Widowiskowo-Sportowej w Katowicach” (Electroacoustic Equipment in the Wojewod Sport Hall in Katowice). *Technika Radia* i Telewizji 4 (1973) S. 1–16.
54. C. McGregor, “Electronic Architecture: the Musical Realities of Its Application,” *6th Int. Conf. Sound Reinforcement*, Nashville, 1988.
55. “Sound System Design for the Eugene Performing Arts Center, Oregon,” *Recording-Eng./Producer*, (1982) S.86–93.
56. P.H. Parkin, and K. Morgan, “‘Assisted Resonance’ in the Royal Festival Hall, London,” *J. Sound and Vibration 2* (1965) 1, S.74–85.
57. “A Special Report on the Experimental (“Assisted Resonance”) System in the Royal Festival Hall, London,” *J. Sound and Vibration*. 1 (1964) 3, S.335–342.
58. N.V. Franssen, “Sur l`amplification des champs acoustiques,” *Acustica 20* (1968), S. 315 ff.

59. M. Ohsmann, “Analyse von Mehrkanalanlagen” (Analyzing multichannel systems), *Acustica* 70 (1990) 4, S. 233–246.
60. S.H. de Koning, “Konzertsäle mit variablem Nachhall,” *Funkschau 57* (1985) 14, S. 33–38.
61. A J. Berkhout, “A Holographic Approach to Acoustic Control,” *J. Audio Eng. Soc. 36* (1988) 12, S. 977–995.
62. P.W. Barnett, “A Review of Reverberation Enhancement Systems,” *AMS Acoustics Ltd.*, London 1988 (U.S. Patent 4.649.564, March 10, 1987).
63. D. Griesinger, “Improving Room Acoustics Through Time-Variant Synthetic Reverberation,” *AES Convention Paris*, Preprint 3014 (1990).
64. D. Griesinger (U.S. Patent 5.109.419, April 28, 1992).
65. W. Prinssen, and B. Kok, “Technical Innovations in the Field of Electronic Modification of Acoustic Spaces,” *Proc. of the IOA*, (1994), Vol. 16, Part 4.
66. Hideo Miyazaki, Takayuki Watanabe, Shinji Kishinaga, and Fukushi Kawakami, Yamaha Corporation, Advanced System Development Center, “Active Field Control (AFC), Reverberation Enhancement System Using Acoustical Feedback Control,” *115th AES Convention*, (October 2003), N.Y.
67. M.A. Poletti, “On Controlling the Apparent Absorption and Volume in Assisted Reverberation Systems,” *Acustica 78*, (1993).
68. W. Ahnert, “Einsatz elektroakustischer Hilfsmittel zur Räumlichkeitssteigerung, Schallverstärkung und Vermeidung der akustischen Rückkopplung,”. Diss. Techn. Univ., Dresden, 1975.
69. J. P. Vian, and X. Meynal, “How Reverberation Enhancement Systems Based on the Virtual Wall Principle can Improve the Acoustic Qualities of Concert Halls from the Point of View of theMusicians, Experiences with the Carmen System,” Symposium on Authentic Music Performance and Variable Acoustics, University of Ghent, May 15, 1998.

Further literature

J. Blauert, *Raeumliches Hoeren (Spatial Hearing)*, Cambridge, MA and London/England: The MIT Press, 1983.
F.A. Everest *The Master Handbook of Acoustics*, 3rd ed., Tab Books, Division of McGraw-Hill, Inc., 1994.
J. Cowan, *Architectural Acoustics Design Guide*. McGraw-Hill Companies, Inc., 2000.
L. Marshall, *Architectural Acoustics*, Elsevier Academic Press, 2006.
T.D. Rossing, *Handbook of Acoustics*, Springer, 2007.

Chapter 8

Stadiums and Outdoor Venues

by Eugene T. Patronis, Jr.

8.1 Overview 203
8.2 Sound Projection Distances 203
8.3 Source Level Requirements 203
8.4 Atmospheric Effects 203
8.5 Techniques for Achieving High Acoustic Pressures 205
8.6 Single Source Location 210
8.7 Covered Seating 210
8.8 Distributed Systems 210
8.9 Private Suite Systems 211
Bibliography. 211

This page intentionally left blank

8.1 Overview

Stadiums and outdoor venues present designers with a set of challenges which are not usually encountered in interior spaces. The leading challenge is the immense distance over which sound of an appreciable level must be projected. This is followed by the fact that the sound is not propagating in a stable medium—i.e., outdoors the air temperature and relative humidity are erratic variables and the air is hardly ever still. Lastly, in addition to the normal 6 dB loss for doubling of distance from a point source in a free field, there exists an additional attenuation from atmospheric absorption whose value is a function of frequency and depends on both temperature and relative humidity. These challenges will be addressed in turn.

8.2 Sound Projection Distances

The values of the distances required between reinforcement loudspeakers and observers in stadiums hinge on both the stadium geometry and on whether the reinforcement system is to be of the single source or distributed loudspeaker type. A distributed system avoids large throw distances and major atmospheric effects but is more expensive to install and maintain. The sound quality of distributed systems in stadiums is somewhat unnatural in that the nonlocal loudspeakers are sources of apparent echoes. A single source system is less expensive to install and maintain but requires special techniques to achieve adequate levels at large throw distances. Advocates exist for both system types. The problems associated with the single source system are more interesting and are the ones initially discussed here.

Throw distances for a central source system in a typical stadium range from 15 m to 200 m (50 ft to 650 ft). Sports stadiums, with the possible exception of baseball stadiums, have playing surfaces in the shape of elongated rectangles with audience seating being peripheral to the playing area. Single source loudspeakers are located at one end of an axis of symmetry along the long dimension of the playing surface as illustrated in Fig. 8-1. This allows the coverage of the seating spaces to fall into a number of zones for which the axial throw distances vary by no more than a factor of two. For example, in the stadium of Fig. 8-1, there is a near, intermediate, and far zone with axial distances of approximately 50 m, 100 m, and 200 m, respectively. Thus, the single source system is actually a splayed array of short, intermediate, and long throw devices.

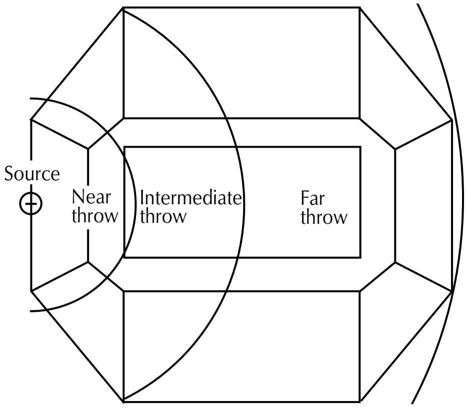


Figure 8-1. Plan view of typical stadium.

8.3 Source Level Requirements

For a point source in a free field without atmospheric absorption, the acoustic pressure varies inversely with the distance measured from the source, i.e., there is a 6 dB loss for each doubling of the distance. The pressure level at 200 meters from such a source is 46 dB less than it is at one meter. If one assumes a noise level of 85 dB and a signal level at least 6 dB above the noise level, then the sound level at one meter must be at least 85 + 6 + 46, or 137 dB, even without any headroom consideration. If one imposes a modest headroom requirement of 6 dB, an impressive 143 dB level is required even before considering atmospheric attenuation.

While this is not attainable from one loudspeaker, it is easily attainable by multiple loudspeakers.

8.4 Atmospheric Effects

Sound propagation is subject to the vagaries of the medium in which it exists. The air in outdoor venues has variable temperature, wind, and relative humidity. The effect of wind is twofold. Wind speed near to the ground is ordinarily less than at a higher elevation. This causes sound waves propagating in a direction into the wind to be diffracted upward while sound waves propagating in the same direction as the wind to be diffracted downward. Cross-winds shift the azimuth of the propagation direction towards that of the wind. Thus wind can cause shifts in apparent loudspeaker aiming points. Additionally, sound will propagate greater distances with the wind than against the wind. A gusting or variable wind introduces a temporal quality to these properties. The effect on a listener is that the sound intensity appears to be modulated as the wind gust rises and falls, a layman's description being "it fades in and out."

Sound speed is influenced by air temperature with higher temperatures corresponding to increased sound speed. This relationship is given by

$$c = 20.06\sqrt{T} \qquad (8\text{-}1)$$

where,

c is the sound speed in meters per second,

T is the absolute temperature in degrees Kelvin.

A fixed air temperature has no influence on propagation direction, but thermal gradients can be a source of further diffraction effects. Normal thermal gradients correspond to a temperature decrease with increasing elevation. Such a condition diffracts sound waves upward such that the apparent direction of propagation is elevated. A temperature inversion gradient has just the opposite effect producing an apparent depressed direction of propagation. The severity of these effects obviously depends on the size of the thermal gradients. Typically encountered stadium situations can result in shifts of 5° or more over a distance of 200 m (650 ft). These effects are illustrated in Fig. 8-2.

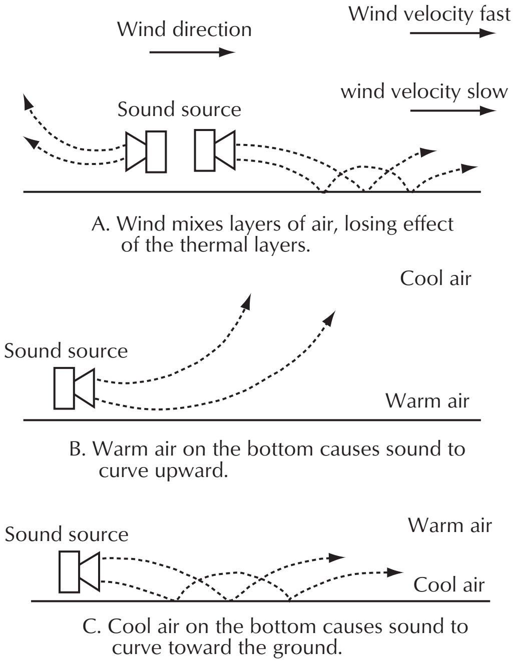

Wind direction
Wind velocity fast
wind velocity slow
Sound source
A. Wind mixes layers of air, losing effect of the thermal layers.
Cool air
Sound source
Warm air
B. Warm air on the bottom causes sound to curve upward.
Sound source
Warm air
Cool air
C. Cool air on the bottom causes sound to curve toward the ground.

Figure 8-2. The effects of wind and thermal gradients on sound propagation.

Atmospheric absorption of acoustical energy ultimately amounts to the conversion of the energy associated with a sound wave into heat energy associated with the random thermal motion of the molecular constituents of the air. Air is basically a gaseous mixture of nitrogen, oxygen, and argon with trace amounts of carbon dioxide, the noble gases, and water vapor. With the exception of argon and the other noble gases all of the constituent molecules are polyatomic and thus have complicated internal structures. There are three mechanisms contributing to the sound energy absorption process. Two of these, viscosity and thermal conductivity, are smooth functions of frequency and constitute what is called the classical absorption. The third or molecular effect involves transfer of acoustic energy into internal energy of vibration and rotation of polyatomic molecules and into the dissociation of molecular clusters. This third effect is by far the most dominant at audio frequencies and explains the complicated influence of water vapor on atmospheric absorption. The detailed behavior given in Fig. 8-3 is illustrative of these effects at a temperature of 20°C whereas the approximate behavior given in Fig. 8-4 is more useful for general calculations.

Table 8-1 is extracted from Fig. 8-3 and illustrates the severity of the absorption effects.

Table 8-1. The Entries Are Attenuation Values in dB/m for Various Values of Relative Humidity at 20°C

RH	0.1 kHz	1 kHz	2 kHz	5 kHz	10 kHz	20 kHz
0%	0.0012	0.0014	0.002	0.0052	0.019	0.07
10%	0.00053	0.018	0.053	0.11	0.13	0.20
100%	0.0003	0.0042	0.010	0.045	0.15	0.50

Below 1 kHz the attenuation is not significant even for a 200 m (650 ft) path length. The relative humidities encountered in practice usually lie in the 10% to 100% range and it can be seen that for frequencies below 5 kHz that wetter air is preferable to drier air. High-frequency equalization to compensate for air losses is usually possible up to about 4 kHz with the amount of equalization required being dependent on the path length. Note that on a dry fall afternoon, the attenuation at 5 kHz over a 200 m path length is about 22 dB. No wonder that a marching brass band on such a day loses its sparkle. As a consequence, long throws in a single source outdoor system are limited to about a 4 kHz bandwidth.

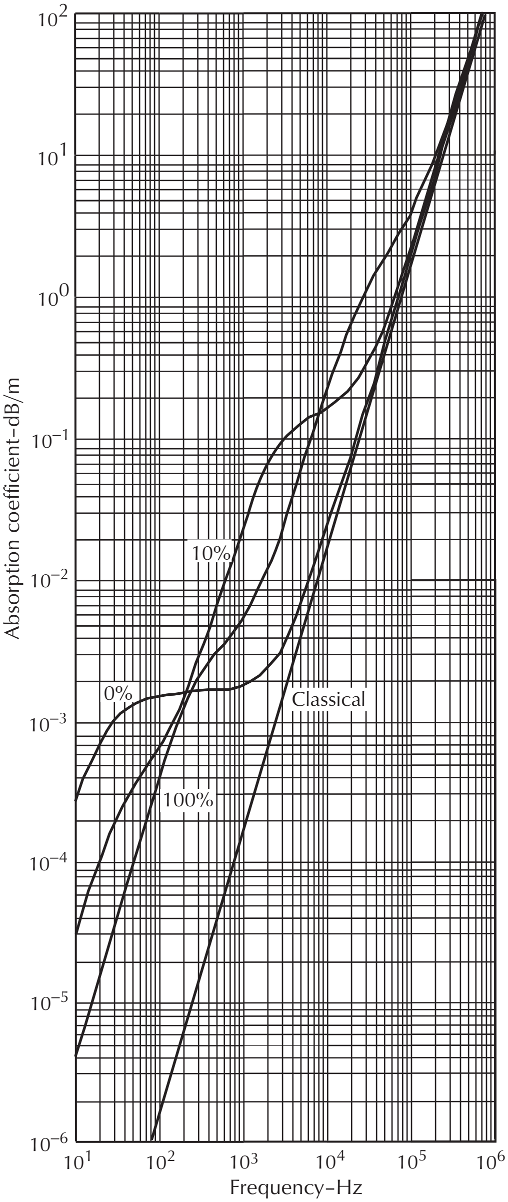


Figure 8-3. Absorption of sound in air at 20°C and one atmosphere for various relative humidities.

8.5 Techniques for Achieving High Acoustic Pressures

In a previous calculation it was shown that for a 200 meter path length, the source must achieve a level of 143 dB at a distance of one meter even in the absence of atmospheric absorption. Including compensation for air losses, the required level can easily rise to 150 dB. Horn throat pressure leading to this level at one meter from the horn mouth would be significantly higher than 150 dB and would suffer from a serious amount of nonlinear distortion. Such pressures are usually achieved by using a coherent array of multiple devices. Typical medium and long throw devices have coverage angles of 40° vertical by 60° horizontal and 20° vertical by 40° horizontal, respectively, these being the angles between the half pressure points of the devices. The long throw angles required in a stadium are usually narrow in the vertical and wide in the horizontal so such devices are stacked to form a vertical array with the axes of the individual devices being parallel. This arrangement for two devices is depicted in Fig. 8-5.

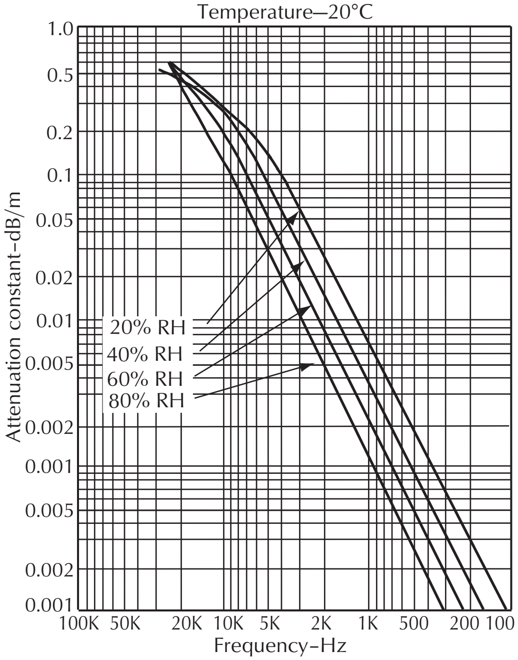


Figure 8-4. Absorption of sound for different frequencies and values of relative humidity.

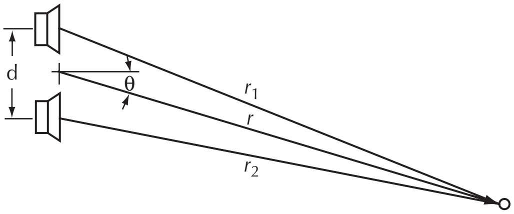


Figure 8-5. A vertical array of two stacked devices

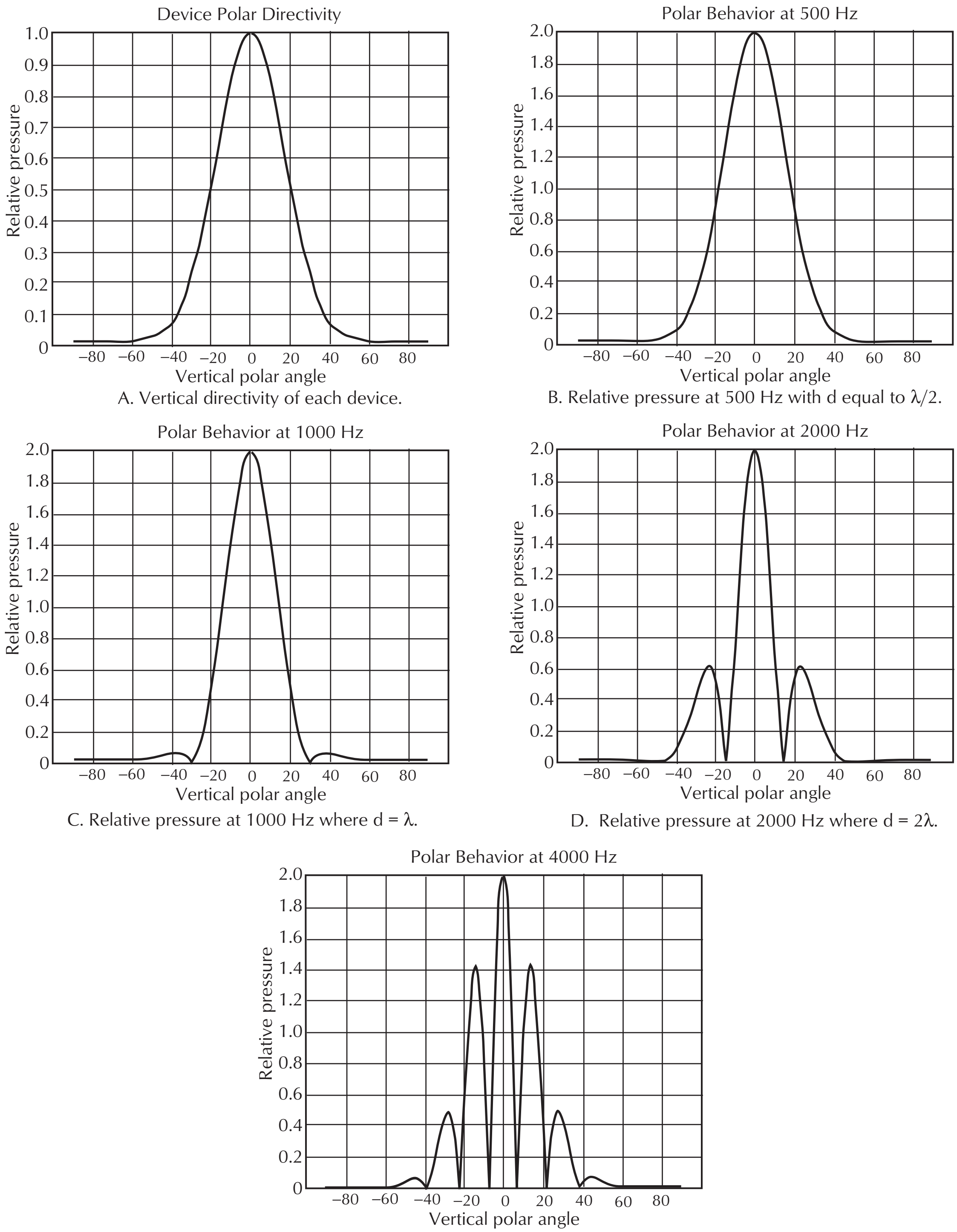
Consider for the moment that the devices are identical point sources that are driven in phase with equal strength electrical signals. In this circumstance, if the observation point, o, is located in the median plane where θ is zero, the acoustic pressure at any radial distance, r, is just double that which would be produced by either source acting alone. This is true because the path lengths are equal so that the two pressure signals undergo the same inverse distance loss and phase lag and thus arrive with equal strength and in phase at the observation point. Now if one considers those observation points where r is always much larger than d and if d is small compared with the wavelength, then the amplitude difference between the two signals as well as the phase difference between the two signals will be insignificant at all such points and again the total pressure will be nearly twice that of a single source acting alone. Such observation points are located in the far field of the combined sources as would be the case in a stadium for all medium or long throw devices. This instance of pressure doubling at all far field observation points only occurs at low frequencies where the wavelength at the operating frequency is significantly larger than the device separation. Consider the case where the operating frequency is such that $d = \lambda/2$. In the far field the amplitude of the signal from each source is again essentially the same but now there will be a phase difference for all values of the angle θ greater than zero. This is most obvious for distant points on the vertical axis where θ = ±90°. At such points the phase difference between the two sources is 180° and the acoustic pressure is zero. The two sources are now exhibiting a frequency-dependent directivity function as a result of their physical placement one above the other. If the two sources are horns rather than point sources, then there is an additional directivity function associated with the horn behavior that is a function of both the azimuthal angle φ as well as the vertical angle θ. The acoustic pressure amplitude for all points in the far field for both sources being driven equally and in phase can be calculated from

$$\begin{aligned} p_m &= \frac{2A}{r}\left|D_h(\theta, \varphi)D_a(\theta)\right| \\ &= \frac{2A}{r}\left|D_h(\theta, \varphi)\cos\left[k\frac{d}{2}\sin(\theta)\right]\right| \end{aligned} \tag{8-2}$$

In the two equations for the pressure amplitude, p_m, A is the source amplitude factor, $D_h(\theta,\varphi)$ is the horn directivity function, and $D_a(\theta)$ is the directivity function brought about by arraying one source above the other. The vertical braces | | denote absolute magnitude of the enclosed quantity. This is necessary as the two directivity functions can independently each be positive or negative dependent upon the frequency and the pressure amplitude is always positive. The quantity, k, is the propagation constant and is related to the wavelength through $k = (2\pi)/\lambda = 2\pi f/c$. It is important to note that the directivity behavior brought on by arraying one device above the other depends only on the angle θ and that the horizontal directivity of the devices in the far field is not influenced by the physical placement of one above the other. This behavior in the vertical plane is depicted in Fig.8-6A through E where the individual devices are 40° vertical by 60° horizontal horns having small mouths.

Figs. 8-6A through E illustrate both the desirable and undesirable attributes of arraying devices in a vertical line. A illustrates the directivity in the vertical plane for each device. The device directivity has a magnitude of 0.5 at ±20° indicating that the vertical coverage angle is 40°. The minimum vertical spacing between the devices is limited by the mouth size and in this instance is 0.344 meter. This corresponds to λ/2 at a frequency of 500 Hz. As shown in Fig. 8-6B, the pressure on-axis is indeed doubled and the overall shape closely follows that of the horn directivity function with just a small narrowing of the vertical coverage angle. In Fig. 8-6C, where the operating frequency is now 1000 Hz, the on-axis pressure is again doubled but now the central lobe is noticeably narrower and small side lobes are in evidence. This trend continues in Fig. 8-6D where the operating frequency is now 2000 Hz. Side lobes are now much in evidence and the central lobe is narrowed even further. Finally, at 4000 Hz, as depicted in Fig. 8-6E, another pair of side lobes appear, the original side lobes, while having narrowed, are considerably stronger, and the central lobe is narrower still while maintaining double on-axis pressure. In all instances the overall envelope containing the vertical directivity behavior of the stacked pair has the same shape as that of the individual device directivity function.

One is not limited to stacking just two devices in a vertical line. Any number, N, of identical devices can be so arranged and when several discrete devices are so arranged the combination is called a *line array.* The qualitative behavior of such an array as observed in the far field is quite similar to that of the stacked pair discussed above. Directional control does not appear until the length of the array becomes comparable to the wavelength, the on-axis pressure in the far field is N times as great as that of a single device, and as the operating frequency increases, side lobes appear and the central lobe becomes narrower and narrower as the operating frequency increases. This assumes that all of



Device Polar Directivity

A. Vertical directivity of each device.

Polar Behavior at 500 Hz

B. Relative pressure at 500 Hz with d equal to λ/2.

Polar Behavior at 1000 Hz

C. Relative pressure at 1000 Hz where d = λ.

Polar Behavior at 2000 Hz

D. Relative pressure at 2000 Hz where d = 2λ.

Polar Behavior at 4000 Hz

E. Relative pressure at 4000 Hz where d = 4λ.

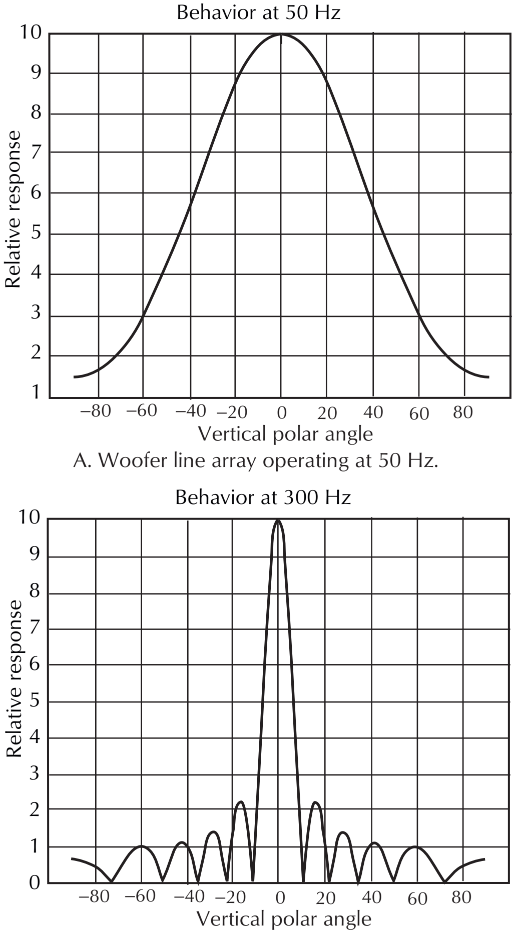
Figure 8-6. Behavior in the vertical plane of stacking two 40° vertical by 60° horizontal small mouth horns.

the devices are identical, have equal amplitude drive signals, and are driven in phase. When driven in this fashion, the array is said to be unprocessed. Modern line arrays are usually structured from enclosures that are each full range loudspeaker systems. Each such loudspeaker system usually divides the audio band into three or four separate bands. In effect then, one is dealing with not just a single line array, but rather three or four arranged parallel to each other. This technique allows optimization in each frequency band with regard to the number of devices, device spacing, and individual device directivity. Fig. 8-7A and B illustrates the performance of the low-frequency section of a straight line array consisting of ten 15 inch diameter woofers with a separation between woofers of 0.6 meter. Fig. 8-7A shows the pressure generated by the array relative to that produced by a single device when operating at 50 Hz, which is the bottom end of the woofer pass band. At this frequency the woofer itself is omnidirectional and the vertical directivity is that produced by the array structure itself. At 300 Hz, which is the upper end of the woofer pass band, the central lobe has been greatly narrowed and there are numerous side lobes as illustrated in Fig. 8-7B.

The operation in the other frequency bands of a full range line array that is unprocessed is qualitatively the same even though the number of devices and spacing between individual devices is, in general, different. The large increase of the pressure on-axis is the desirable attribute whereas the accompanying narrowing of the central lobe and the generation of side lobes are undesirable. The latter behavior can be mitigated to some extent by arraying on an arc rather than a straight line. The mounting hardware linking the devices in an array is structured so as to allow a splay between individual units with an adjustable angle in the range of 2° to 5°. This shapes the array into an arc of a circle rather than in a straight line. In such an arrangement, there is some reduction in the maximum pressure on axis, but the central lobe retains a more uniform width particularly at the upper ends of the various frequency bands. Mathematical details may be found in the first reference at the end of this chapter.

Another arraying technique worthy of mention is that of the Bessel array first introduced by the Dutch industrial giant Philips. The Bessel array in its simplest configuration employs five identical elements and although it only doubles the on-axis pressure in the far field, it does so while having a coverage pattern both vertically and horizontally that matches the coverage pattern of the individual elements from which it is constructed. The individual elements might be woofers, horns, or full range systems of any type. In the simplest configuration, five identical devices are arrayed along a straight line either horizontally or vertically as close together as possible with parallel axes. The unique properties of the array are brought about by weighting the voltage drive to the array in the sequence 0.5, 1,1, –1, and 0.5. For example for a vertical array, one-half of the available voltage drive is applied to the top and bottom elements. This is easily accomplished in practice by connecting these two elements in series with each other. The interior elements are then connected together in parallel with the lowest interior element being operated in reverse polarity. This physical and electrical arrangement appears in Fig. 8-8.

When observed in the far field this arrangement produces twice the pressure of that of a single device



A. Woofer line array operating at 50 Hz.

B. Woofer line array operating at 300 Hz.

Figure 8-7. Woofer line array operating at 50 Hz and 300 Hz.

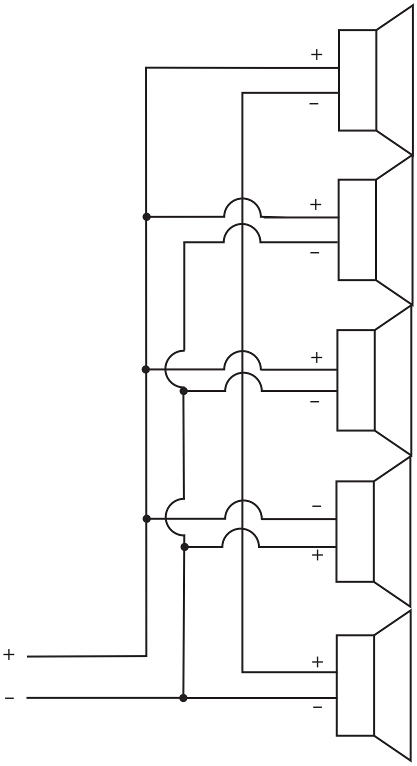


Figure 8-8. Physical and electrical arrangement of simplest Bessel array.

and has a directivity that is almost identical to that of a single structural element. The on-axis amplitude and phase response is also that of a single constituent device. For observation points off-axis, where θ is no longer zero, the amplitude response will exhibit a small ripple as the frequency changes throughout the pass band of the devices employed. The magnitude of this ripple is inconsequential as it amounts to less than 1.25 dB. Perhaps of more importance is the behavior of the phase response when observed off-axis. The phase response ripples between plus and minus 90° as the frequency changes throughout the pass band of the devices that are employed. This ripple in phase response is superimposed on the normal phase response of an individual device. The mathematical details describing the Bessel array behavior are given in the first reference given at the end of this chapter. The Bessel array, having preserved the coverage pattern of a single device, is relatively insensitive to aiming errors resulting from wind or thermal effects.

In recent years Meyer Sound Laboratories, Inc. has produced a number of well-designed self-powered loudspeaker systems that are well adapted for employment in stadiums and outdoor venues. Even though the concept of self-powered loudspeakers is an old one, Meyer took the concept several technical steps further by including not only the appropriate power amplification but also the necessary signal processing, amplifier, and loudspeaker protection circuitry as well, all within the confines of the loudspeaker enclosure. One such system that can be a real problem solver is Meyer's SB-1 sound beam loudspeaker system that is depicted in Fig. 8-9.

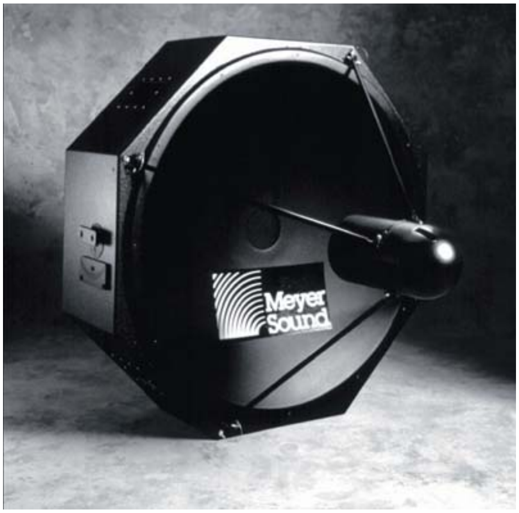


Figure 8-9. Meyer SB-1 Sound Beam. Used with the permission of Meyer Sound Laboratories, Inc.

The system of Fig. 8-9 is based on the properties of a parabolic reflector. A parabolic reflector is really a paraboloid, which is the shape generated when a parabola is rotated about its principal axis. Such shapes have been employed for many years as the basis for reflecting telescopes, radar antennas, and microphones. In the case of a telescope or microphone application the paraboloid focuses a beam of light or sound emanating from great distances to a common point known as the focal point of the paraboloid. In the instance of a radar antenna that is employed for both transmission and reception, radiation from a small element located at the focal point is formed into a parallel beam having small divergence during transmission and return signals traveling parallel to the system's axis are focused on the small element at the focal point during reception. In the sound beam application of the SB-1 a 4 in compression driver fitted with an aspherical horn is mounted in a bullet-shaped pod and located at the focal point of the paraboloid. This assembly is aimed at a fiberglass parabolic reflector having a diameter of approximately

50 in. In order for such a reflector to form a well-defined beam it is necessary that the diameter of the reflector be considerably larger than the wavelength. This is true for this unit except at its low-frequency limit of 500 Hz. In the vicinity of 500 Hz diffraction produces some out-of-beam energy that is undesirable. This is compensated for through the placement of a 12 in diameter cone driver within the enclosure behind the reflector. This driver radiates through an opening at the center of the reflector. This driver's electrical signal is band-pass limited in the vicinity of 500 Hz and is phased to cancel the diffracted out-of-beam signal produced by the compression driver. Separate power amplifier and processing circuitry is provided for the two drivers with all of the electronics and associated power supplies being located in the main enclosure. Once the system is assembled and aimed it is necessary only to supply the appropriate ac power and audio signal. The system is specified to produce a maximum SPL of 110 dB at 100 m in a pass band of 500 to 15,000 Hz with a coverage angle of 20°.

8.6 Single Source Location

Single source loudspeaker arrays are usually mounted at one end of an axis of symmetry of the stadium seating with the long axis being preferred. It is desirable to place the cluster at such an elevation that the components of the array are aimed down on the audience. This positioning minimizes the spill of sound into the surrounding community.

8.7 Covered Seating

Many stadiums feature double and occasionally triple decking such that a portion of the lower seating is obscured from a line-of-sight view of the single source point. In this instance the perception of a single source can be maintained while still providing direct sound to the covered seats by creating a stepped zone delay system. In this system, a distributed loudspeaker system is installed beneath the upper deck and arranged in a series of coverage zones such that the obscured seats in a given zone are all approximately the same distance from the single source point. The electrical signals to the loudspeakers in a given zone are delayed by an amount equal to the transit time of sound from the single source point to the given zone. If the zone loudspeakers radiate principally in the direction which would have been taken by the single source system had it not been obscured, one generates a traveling source of sound from one zone to the next that is in sync with sound from the single source system. Zonal boundaries at a linear spacing of about 20 m (65 ft) have been found to produce very intelligible apparently echo-free results.

8.8 Distributed Systems

Distributed systems are capable of producing full bandwidth sound throughout a stadium provided that the individual loudspeaker systems are installed with sufficient density such that the axial throw of any given unit is 50 m (165 ft) or less. A throw of 50 m will require variable equalization for air absorption if proper high-frequency balance is to be maintained. The uniformity of sound distribution is improved with an increasing loudspeaker density and hence an increasing expense.

The design of the individual sources in such a system is carried out as one would for an interior space that has a designated seating area. An array of loudspeakers is formed using conventional arraying techniques with the view toward providing uniformity of coverage and full bandwidth. The successive distributed areas are chosen such that the areas overlap at the −6 dB point of an individual area. This will provide quite uniform coverage throughout all of the seating areas. Psychoacoustically, the sound will appear more natural to the listeners if the source is elevated and in front of the audience. Weather proofing techniques must be employed in the loudspeaker manufacture and/or in the loudspeaker installation process.

A distributed system may be powered in a number of ways. All of the power amplifiers may be located at a single central point, in which case, long cable runs must be made on 70 V or 200 V lines to the distributed loudspeakers. This is convenient from the standpoint of monitoring or servicing the amplifiers, but is enormously expensive to install. Rather than locate all amplification at a single position, power amplifiers may be located at several subpoints throughout the stadium. Less costly low-level signal wiring connects the subpoints and high-level power runs are shortened and hence made less expensive. Alternatively, powered loudspeakers are available from which to construct the individual loudspeaker clusters. Less costly low-level signal wiring can now be run to each cluster. Ac power must be made available at each loudspeaker location under this option. This expense, however, is now shifted to the electrical contractor. This individually powered option is the least expensive initially but may present a servicing nightmare in the future.

Regardless of the technique employed for installation, any reasonable design will include provisions for monitoring the operation of the individual loudspeaker systems from a central point. The better, more sophisticated designs will also provide for individual system adjustments from the central monitoring point.

In summary, distributed systems are more expensive to install and considerably more expensive to maintain. They are capable of wider bandwidth than are single source systems and are less sensitive to atmospheric effects. Some listeners object to the apparent echoes produced by distributed systems whereas others maintain that is the way stadiums should sound.

8.9 Private Suite Systems

Practically all new stadium construction as well as renovation incorporate private suites in the stadium concept. These suites provide a view of the playing area through a glass wall but otherwise offer an isolated environment with spaces for seating, dining, and other entertainment. They are provided with complete electronic entertainment and communication systems including closed circuit television for instant replay and pre- and postgame shows. Aside from the installed entertainment system, it is customary to provide a crowd noise feed, a public address feed, and a press feed into each suite with the selection of the active feed or feeds being under the control of the suite occupants. The quality of the installed electronics often surpasses that which is usually associated with a home entertainment system.

Bibliography

Don Davis and Eugene Patronis, Jr., *Sound System Engineering*, 3rd edition, Burlington: Focal Press, 2006.
Kinsler, Frey, Coppens, and Sanders, *Fundamentals of Acoustics*, 4th edition, New York: John Wiley, 2000.

This page intentionally left blank

Chapter 9

Modeling and Auralization

by Dominique J. Chéenne, Ph.D.

9.1 Introduction 215
9.2 Acoustical Modeling 215
  9.2.1 Physical Models 215
    9.2.1.1 Frequency and Wavelength Considerations in Physical Models 216
    9.2.1.2 Time and Distance Considerations in Physical Models 217
    9.2.1.3 Medium Considerations in Physical Models 217
    9.2.1.4 Source and Receiver Considerations in Physical Models 218
    9.2.1.5 Surface Materials and Absorption Considerations in Physical Models 219
  9.2.2 Computational Models 220
    9.2.2.1 Geometrical Models 220
    9.2.2.2 Wave Equation Models 225
9.2.2.2.1 Analytical Model: Full-Wave Methodology. 226
    9.2.2.3 Numerical Model: Boundary Element Methodology 227
    9.2.2.4 Statistical Models 228
    9.2.2.5 Small Room Models 228
  9.2.3 Empirical Models 229
    9.2.3.1 Gypsum Cavity Wall Absorption 229
    9.2.3.2 Absorption from Trees and Shrubs 230
  9.2.4 Hybrid Models. 230
9.3 Auralization 231
  9.3.1 The Basic Auralization Process. 232
  9.3.2 Implementation 232
    9.3.2.1 Binaural Reproduction Using Loudspeakers 232
    9.3.2.2 Binaural Reproduction Using Headphones 232
    9.3.2.3 Multichannel Reproduction Using Loudspeakers 233
  9.3.3 Real-Time Auralization and Virtual Reality 233
9.4 Conclusion 235
References 236

This page intentionally left blank

9.1 Introduction

As it is often the case with branches of engineering dealing with the understanding and the prediction of complex physical phenomena, *modeling* has rapidly become an integral part of the acoustical design process. When dealing with indoor sound propagation the use of a suitable model may allow the designer to assess the consequences that a change in parameters such as room shape, material selection, or source placement will have on variables like sound pressure, reverberation time, or reflection ratios at specific points inside the room. Acoustical models can also be developed for outdoor sound studies as one may inquire what shape and height a highway barrier needs to be in order to attenuate a specific level of unwanted noise from highway traffic over a given area. In these instances, the model is expected to provide the designer with the answer to the fundamental *“what if?”* question that is at the genesis of an engineered design—i.e., one that leaves little to chance in terms of the sensible selection of its parameters in order to achieve a specific result. The answers to the question allow the designer to assess the performance or the cost-effectiveness of a design, based on a specific set of criteria prior to committing to it.

Although an experienced designer may be able to achieve a substantial understanding of the acoustics of a given environment simply by looking at the data that results from the modeling phase of the design, the acoustical model may also be used to provide an auditory picture of the data so that a qualitative evaluation of the acoustics can be performed by trained and/or untrained listeners. This phase of the design—known as *auralization*—aims at doing for one’s ears what a picture does for one’s eyes: present a description of an environment in a way that is best suited to the most appropriate sensor. In this instance, the basic goal is to use sound to demonstrate what a specific environment will sound like, just like the fundamental purpose of an image is to illustrate what an environment may look like. The challenges associated with virtual representation that exist in the visual world, such as accuracy, context, and perception, are also present in the aural environment, and the old engineering school adage that “a picture is worth a thousand words, but only if it is a good picture” rings also true in the world of acoustical modeling and of auralization.

The aim of this chapter is to provide the reader with a basic understanding of the various methodologies that are used in acoustical modeling and auralization, and the emphasis is on models that can be used for the evaluation of room acoustics. The reader is referred to the bibliography for further in-depth reading on the topic.

9.2 Acoustical Modeling

This section will review various acoustical modeling techniques from the perspective of theory, implementation, and usage.* The categorization and grouping of the modeling techniques into three general families (physical models, computational models, and empirical models) and in further subgroups as presented in Fig. 9-1 is provided as a means to identify the specific issues associated with each technique in a fashion that is deemed effective by the author from the standpoint of clarity. A brief mention of hybrid models that combine various techniques will also be introduced. The sections of this chapter are written as independently from each other as possible so the reader can go to topics of specific interest without reading the preceding sections.

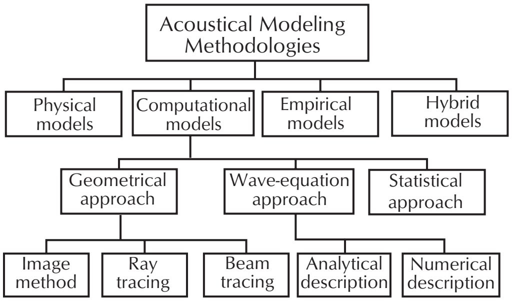


Figure 9-1. General classification of acoustical modeling methodologies.

9.2.1 Physical Models

This class of model uses a scaling approach to yield a 3D representation of typically large acoustical environments like theater or auditoria for the purpose of evaluation and testing. The models are constructed at geometrical scales ranging from 1:8 to 1:40, and an example of a 1:20 physical model is presented in Figs. 9-2 and 9-3. Physical modeling techniques became pop-

---

[*] Model (mäd’l) *n.* [Fr. *modèle* < It *modello*] **1**. a) a small copy or imitation of an existing object made to scale b) a preliminary representation of something, serving as the plan from which the final, usually larger, object is to be constructed c) archetype d) a hypothetical or stylized representation e) a generalized, hypothetical description, often based on an analogy, used in analyzing or explaining something.

From *Webster New World College Dictionary.* 4th Edition (2000) ular after WWII and are still used today by some designers when access to other modeling tools is limited and/or when a physical representation of the space under review is needed for the purpose of testing as well as visualization.

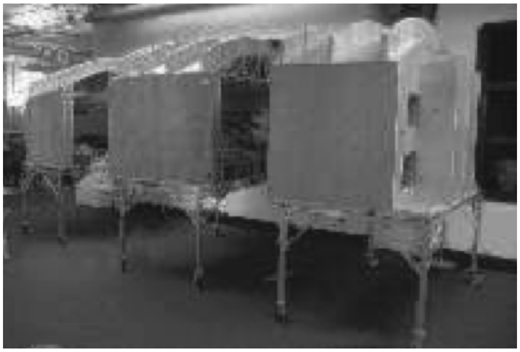


Figure 9-2. External view of a 1:20 scale physical model. Courtesy Kirkegaard Associates.

The scaled approach used in the geometrical construction of the model implies that in order to investigate some of the relevant physical parameters,[1] specific acoustic variables will also need to be scaled accordingly.

9.2.1.1 Frequency and Wavelength Considerations in Physical Models

The issues pertaining to frequency and wavelength in scale models are best presented with the use of an example. If a source of sound generating a frequency $f = 1000$ Hz is placed inside a room, then under standard conditions of temperature ($t = 20°C$) where the velocity of sound is found to be $c = 344$ m/s, the wavelength of the sound wave is obtained from

$$\lambda = \frac{c}{f} \tag{9-1}$$

or in this example

$$l = 0.34 \text{ m}$$

The wave number k can also be used to represent the wavelength since

$$k = \frac{2\pi}{\lambda} \tag{9-2}$$

so in our example we have

$$k = 18.3 \text{ m}^{-1}$$

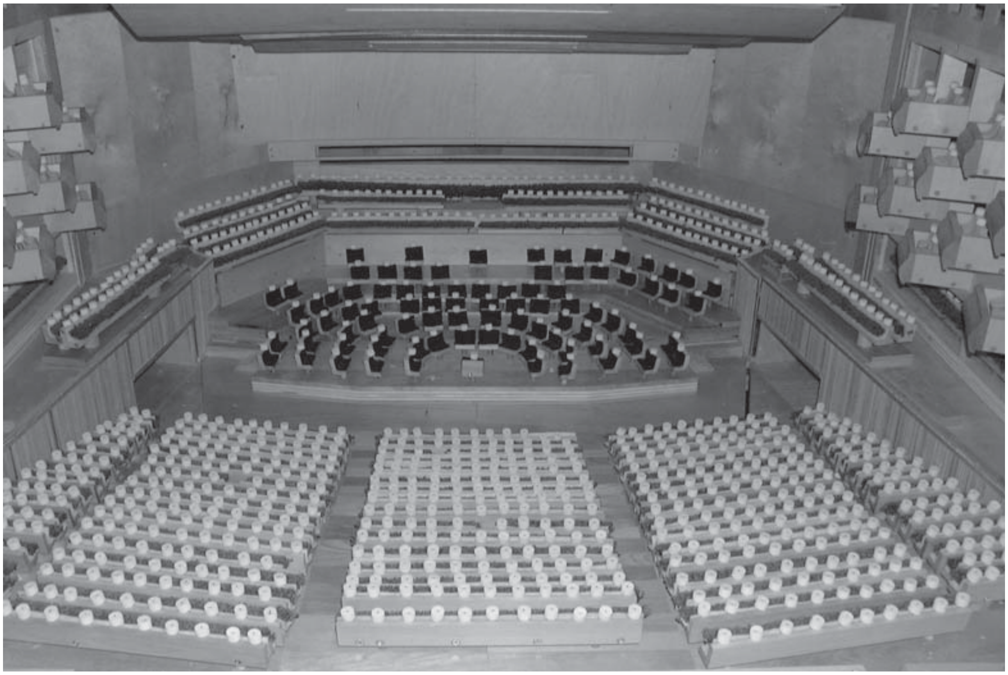


Figure 9-3. Internal view of a 1:20 scale physical model. Courtesy Kirkegaard Associates.

In the absence of acoustical absorption, the *relative* size of the wavelength of a sound wave to the size of an object in the path of the wave dictates the *primary* physical phenomenon that takes place when the sound waves reach the object. If the object can be considered acoustically hard (i.e., having a low absorption coefficient) so that the energy of the wave is not significantly reduced by absorption and if the characteristic dimension of the object (its largest dimension in the path of the sound wave) is denoted by a then the product ka can be used to predict how the sound waves will be affected by the presence of the object.

Table 9-1 shows a range of products ka and the resulting primary effect that the object will have on the sound waves.

Table 9-1. Effect of Wave Number and Object Dimension on the Propagation of Sound Waves

Value of *ka*	Primary Phenomenon Taking Place When the Sound Waves Reach the Object (Not Including Absorption)
$ka \le 1$	*Diffraction*: the sound waves travel around the object without being affected by its presence. The object can be considered invisible to the waves.
$1 < ka < 5$	*Scattering*: the sound waves are partially reflected by the object in many directions and in a complicated fashion. This scattering phenomenon is associated with the notion of acoustical diffusion.
$ka > 5$	*Reflection*: the sound waves are deflected by the object in one or more specific direction(s) that can be predicted from application of basic geometry laws.

In our example if an acoustically hard object of dimension $a = 0.5$ m is placed inside a full-size room, then the sound waves will be reflected by the object since $ka = 9.1$, a value clearly above the lower limit of 5 at which reflections become the primary phenomenon.

If a 1:10 scale model of the room is now created and the object is scaled by the same amount its dimension has now become $a' = 0.05$ m. Under the earlier conditions where $f = 1000$ Hz the product ka' now has a value of $ka' = 0.91$ and the conclusion that must be drawn from the guidelines presented in Table 9-1 is that the sound waves diffract around the object: in other words, the model has failed at predicting accurately the primary physical phenomenon of sound reflection that takes place when a 1000 Hz sound wave strikes an acoustically hard object of 0.5 m in size.

In order for the model to yield the proper conclusion—i.e., the sound waves are reflected by the object—the wavelength of the sound waves has to be scaled down by the same amount as the physical dimensions of the room, or looking at Eq. 9-1—and keeping the speed of sound a constant—the frequencies of the sound waves have to be scaled up by an amount equal to the inverse of the model's physical scale. In our example, a 10 kHz frequency needs to be used in the model in order to assess the conditions that exist inside the full-size room under 1 kHz excitation. If data pertaining to the room needs to be available from 50 Hz to 20 kHz, the use of a 1:10 scale physical model will require that frequencies from 500 Hz to 200 kHz be used during the investigation.

9.2.1.2 Time and Distance Considerations in Physical Models

A sound wave traveling at a velocity c will take a time t to cover a distance x according to the relation

$$x = ct \tag{9-3}$$

In a physical model the dimensions are scaled down and as a consequence the time of travel of the waves inside the model is reduced by the same factor. If time-domain information with a specific resolution is required from the model, then the required resolution in the time data must increase by a factor equal to the inverse of the scale in order to yield the desired accuracy.

As an example, if a sound source is placed at the end of a room with a length $x = 30$ m and $c = 344$ m/s, the application of relation, Eq. 9-3, shows that the sound waves can be expected to reach the other end of the room in 87.2 ms. If an accuracy of ±10 cm is desired in the distance information, then a time resolution of ±291 μs is required for the time measurements.

In a 1:10 scale situation—and under the same conditions of sound velocity—the sound waves will now take 8.72 ms to travel the length of the model and a resolution of ±29.1 μs will be required in the time-measuring apparatus to yield the required distance resolution.

9.2.1.3 Medium Considerations in Physical Models

As a sound wave travels through a gaseous medium like air it loses energy because of interaction with the molecules of the media in a phenomenon known as *thermal relaxation*. Energy is also lost via spreading of the wave as it travels away from the source. The absorption in the medium as a function of distance of travel x and other parameters such as temperature and humidity can be represented by a loss factor K that is given by:

$$K = e^{-mx} \tag{9-4}$$

where,

m is called the *decay index*, and it takes into account the temperature and the humidity of the air as a function of the frequency of the sound.

The value of m has been determined both analytically and experimentally for various conditions of temperature and humidity over a range of frequencies extending from 100 Hz to 100 kHz and an examination of the data shows that the absorption of air increases with increased frequencies and that for a given frequency the maximum absorption takes place for higher relative humidity.

Since the distance x traveled by sound waves in a physical model are scaled down in a linear fashion (i.e., by the scale factor), one cannot expect the attenuation of the waves inside the model to accurately reflect the absorption of the air since the loss factor K follows an exponential decay that is dependent upon the term m that is itself affected by the physical properties of the medium. In a scaled physical model this discrepancy is taken into account by either totally drying the air inside the model or achieving conditions of 100% relative humidity; in either case, the approach yields a simpler relation for m that becomes solely dependent on temperature and frequency. For example, in the case of totally dry air the decay index becomes

$$m = (33 + 0.2T)10^{-12}f^2 \tag{9-5}$$

and under these conditions, it is clear that the dominant term that dictates the loss of energy in the sound wave is the frequency f since m is proportional to f^2 and varies only slightly with the temperature T.

Another available option to account for the differences in the air absorption between a scaled physical model and its full-size representation is to use a different transmission medium in the model. Replacing the air inside the model with a simple molecular gas like nitrogen will yield a decay index similar to that of Eq. 9-5 up to frequencies of 100 kHz but this is a cumbersome technique that limits the usability of the scale model.

9.2.1.4 Source and Receiver Considerations in Physical Models

To account for the primary phenomena (defined in Table 9-1) that take place over a range of frequencies from 40 Hz to 15 kHz inside a full-size room, one needs to generate acoustic waves with frequencies extending from 400 Hz to 150 kHz if a 1:10 model is used, and the required range of test frequencies becomes 800 Hz to 300 kHz in the case of a 1:20 scale model. The difficulties associated with creating efficient and linear transducers of acoustical energy over such frequency ranges are a major issue associated with the use of physical scale models in acoustics.

Transducers of acoustical energy that can generate continuous or steady-state waves over the desired range of frequencies and that can also radiate the acoustical energy in a point-source fashion are difficult to design, thus physical scale models often use impulse sources for excitation; in these instances the frequency information is derived from application of transform functions to the time-domain results. One commonly used source of impulse is a spark generator as shown in Fig. 9-4 where a high voltage of short duration (typically ranging from less than 20 μs to 150 μs) is applied across two conductors separated by a short distance. A spark bridges the air gap and the resulting noise contains substantial high-frequency energy that radiates over an adequately spherical pattern.

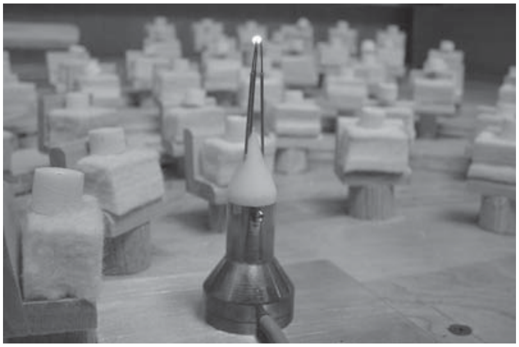


Figure 9-4. A spark generator used in scale physical models. Courtesy Kirkegaard Associates.

Although the typical spark impulse may contain sufficient energy beyond 30 kHz, the frequency response of a spark generator is far from being regular and narrowing of the bandwidth of the received data is required in order to yield the most useful information. The bandwidth of the impulse $\Delta_{f\,impulse}$ and its duration $\tau_{impulse}$ are related by the uncertainty principle

$$\Delta f_{impulse} t_{impulse} = 1 \tag{9-6}$$

When dealing with impulses created with a spark generator, the received data must be processed via a bandpass filter of order to eliminate distortion associated with the nonlinearity of the spark explosion, but the filter must have a sufficient bandwidth $\Delta f_{filter\ r}$ to avoid limiting the response of the impulse.

$$\Delta f_{filter} \geq \frac{4}{t_{impulse}} \tag{9-7}$$

Steady-state sound waves can be generated over spherical (nondirectional) patterns up to frequencies of about 30 kHz by using specially designed electrostatic transducers. Gas nozzles can also be used to generate continuous high-frequency spectra although issues of linearity and perturbation of the medium must be taken into account for receiver locations that are close to the source.

Microphones are available with very good flatness (±2 dB) over frequency responses extending beyond 50 kHz. The main issues associated with the use of microphones in physical scale models are that the size of the microphone cannot be ignored when compared to the wavelength of the sound waves present in the model, and that microphones become directional at high frequencies. A typical half-inch microphone capsule with a 20 cm housing in a 1:20 scale model is equivalent to a 25 cm × 4 m obstacle in a real room and can hardly be ignored in term of its contribution to the measurement; furthermore its directivity can be expected to deviate substantially (>6 dB) from the idealized spherical pattern above 20 kHz. Using smaller capsules (¼ in or even ⅛ in) can improve the omnidirectivity of the microphone but it also reduces its sensitivity and yields a lower SNR during the measurements.

9.2.1.5 Surface Materials and Absorption Considerations in Physical Models

Ideally, the surface materials used in a scale physical model should have absorption coefficients that closely match those of real materials planned for the full-size environment at the equivalent frequencies. For example, if a 1:20 scale model is used to investigate sound absorption from a surface at 1 kHz in the model (or 50 Hz in the real room) then the absorption coefficient a of the material used in the model at 1 kHz should match that of the planned full-size material at 50 Hz. In practice this requirement is never met since materials that have similar absorption coefficients over an extended range of frequencies are usually limited to hard reflectors where $a < 0.02$ and even under these condition, the absorption in the model will increase with frequency and deviate substantially from the desired value. The *minimum* value for the absorption coefficient of any surface in a model can be found from

$$\alpha_{min} = 1.8 \times 10^{-4} \sqrt{f} \tag{9-8}$$

where,

f is the frequency of the sound wave at which the absorption is measured.

Thus at frequencies of 100 kHz, an acoustically hard surface like glass in a 1:20 scale model will have an absorption coefficient of $\alpha_{min} = 0.06$, a value that is clearly greater than $\alpha < 0.03$ or what can be expected of glass at the corresponding 5 kHz frequency in the full-size space.

The difference in level between the energy of the *n*th reflected wave to that of the direct wave after *n* reflections on surfaces with an absorption coefficient a is given by

$$\Delta_{Level} = 10\log(1-\alpha)^{n} \tag{9-9}$$

Considering glass wherein the model $\alpha = \alpha_{min} = 0.06$, the application of Eq. 9-9 above shows that after two reflections the energy of the wave will have dropped by 0.54 dB. If the reflection coefficient is now changed to $\alpha = 0.03$ then the reduction in level is 0.26 dB or a relative error of less than 0.3 dB. Even after five reflections, the relative error due to the discrepancies between a and a_{min} is still less than 0.7 dB, a very small amount indeed.

On the other hand, in the case of acoustically absorptive materials ($\alpha > 0.1$) the issue of closely matching the absorption coefficients in the models to those used in the real environment becomes very important. The application of Eq. 9-9 to absorption coefficients α in excess of 0.6 shows that even a slight mismatch of 10% in the absorption coefficients can result in differences of 1.5 dB after only two reflections. If the mismatch is increased to 20% then errors in the predicted level in excess of 10 dB can take place in the model.

Due to the difficulty in finding materials that are suitable for use in both the scaled physical model and in the real-size environment, different materials are used to match the absorption coefficient in the model (at the scaled frequencies) to that of the real-size environment at the expected frequencies. For example, a 10 mm layer of wool in a 1:20 scale model can be used to model rows of seats in the actual room, or a thin layer of polyurethane foam in a 1:10 scale model can be used to represent a 50 mm coating of acoustical plaster in the real space. Another physical parameter that is difficult to account for in scale physical model is stiffness*,* thus the evaluation of effects such as diaphragmatic absorption and associated construction techniques is difficult to model accurately.

9.2.2 Computational Models

This section presents models that create a mathematical representation of an acoustical environment by using assumptions that are based on either a geometrical, analytical, numerical, or statistical description of the physical phenomena (or parameters) to be considered, or on any combination of the afore- mentioned techniques, In all instances, the final output of the modeling phase is the result of extensive mathematical operations that are usually performed by computers. With the development of powerful and affordable computers and of graphical interfaces these modeling tools have become increasingly popular with acoustical designers. To various degrees, the aim of computational models is to ultimately yield a form of the impulse response of the room at a specific receiver location from which data pertaining to time, frequency, and direction of the sound energy reaching the receiver can be derived. This information can then be used to yield specific quantifiers such as reverberation time, lateral reflection ratios, intelligibility, and so on.

The inherent advantage of computational models is flexibility: changes to variables can be made very rapidly and the effects of the changes are available at no hard cost, save for that of computer time. Issues related to source or receiver placement, changes in materials and/or in room geometry can be analyzed to an infinite extent. Another advantage of computational models is that scaling is not an issue since the models exist in a virtual world as opposed to a physical one.

Computational models are themselves divided into subgroups that are fundamentally based on issues of adequacy, accuracy, and efficiency. An *adequate* model uses a set of assumptions based on a valid (true) description of the physical reality that is to be modeled. An accurate model will further the cause of adequacy by providing data that is eminently useful because of the high confidence associated with it. An *efficient* model will aim at providing fast and adequate results but maybe to a lesser—yet justified—extent in accuracy. Although issues of accuracy and of efficiency will be considered in this portion of the chapter, the discussion of the various classes of computational models will be primarily based on their adequacy.

9.2.2.1 Geometrical Models

The primary assumption that is being made in all geometrical models[2] applied to acoustics is that the wave can be seen as propagating in one or more specific directions, and that its reflection(s) as it strikes a surface is (are) also predictable in terms of direction; this is a very valid assumption when the wavelength can be considered small compared to the size of the surface and the condition $ka > 5$ presented in Table 9-1 quantifies the limit above which the assumption is valid. Under this condition, the propagating sound waves can be represented as straight lines emanating from the sources and striking the surfaces (or objects) in the room at specific points. The laws of optics involving angles of incidence and angles of reflection will apply and the term *geometrical acoustics* is used to describe the modeling technique.

The second assumption of relevance to geometrical acoustics models is that the wavelength of the sound waves impinging on the surfaces must be large compared to the irregularities in the surface, in other words the surface has to appear smooth to the wave and in this instance the irregularities in the surface will become invisible since the wave will diffract around them. If the characteristic dimension of the irregularities is denoted by b, then the condition $kb < 1$ is required using the criteria outlined in Table 9-1. This is a necessary condition to assume that the reflection is *specular*, that is that all of its energy is concentrated in the new direction of propagation. Unless this condition is met in the actual room the energy of the reflected wave will be spread out in a diffuse fashion and the geometrical acoustics assumption of the model will rapidly become invalid, especially if many reflections are to be considered.

Image Models. In this class of geometrical acoustics models, the assumption that is being made is that the only sound reflections that the model should be concerned about are those reaching the receiver, so the methodology aims at computing such reflections within time and order constraints selected by the user of the model while ignoring the reflections that will not reach the receiver. To find the path of a first-order reflection a source of sound S_0 is assumed to have an image—a virtual source S_1—located across the surface upon which the sound waves are impinging as presented in Fig. 9-5.

As long as the surface can be considered to be rigid, the image method allows for the prediction of the angles of reflections from the surface and can find all of the paths that may exist between a source and a receiver.[3] It also satisfies the *boundary conditions* that must take place at the surface, that is, the acoustical pressures have to be equal on both sides of the surface at the reflection point, and the velocity of the wave has to be zero at the interface. The image from the virtual source S_1 can also be used to determine where the second-order reflections from a second surface will be directed to since as far as the second surface is concerned, the wave that is impinging upon it emanated from S_1. A second order source S_2 can thus be created as shown in Fig. 9-6 and the process can be repeated as needed to investigate any order of reflections that constitutes a path between source and receiver.

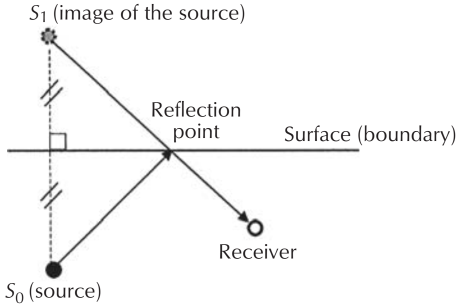

S_1 (image of the source)
Reflection point
Surface (boundary)
Receiver
S_0 (source)

Figure 9-5. A source and its virtual image located across a boundary define the direction of the first-order reflection.

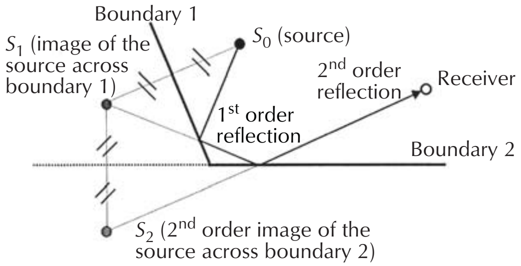

Boundary 1
S_1 (image of the source across boundary 1)
S_0 (source)
2nd order reflection
Receiver
1st order reflection
Boundary 2
S_2 (2nd order image of the source across boundary 2)

Figure 9-6. Higher-order reflections can be created by adding virtual images of the source.

It is thus possible to collect the amplitude and the direction of *all* of the reflections at a specific location as well as a map of where the reflections emanate from. Even the reflection paths from curved surfaces can be modeled by using tangential planes as shown in Fig. 9-7 and Fig. 9-8.

Since the speed of sound can be assumed to be a constant inside the room, the distance information pertaining to the travel of the reflections can be translated into time-domain information; the result is called a *reflectogram* (or sometimes an *echogram*) and it provides for a very detailed investigation of the reflections inside a space that reach a receiver at a specific location. A sample of a reflectogram is shown in Fig. 9-9..

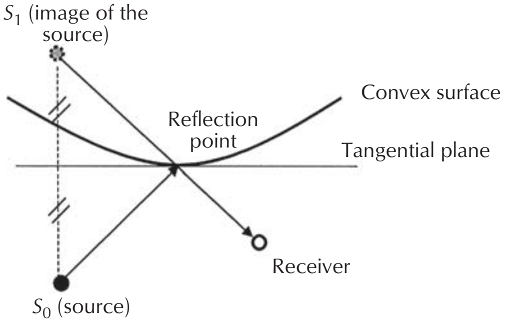

S_1 (image of the source)
Convex surface
Reflection point
Tangential plane
Receiver
S_0 (source)

Figure 9-7. Image construction from a convex plane. Adapted from Reference 1.

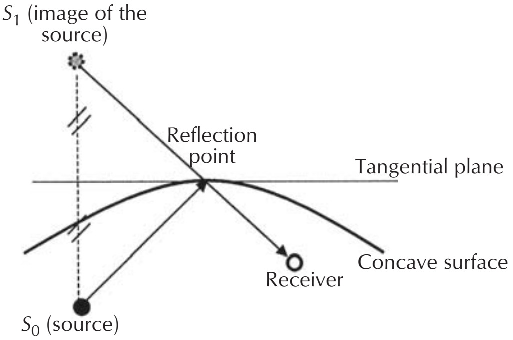

S_1 (image of the source)
Reflection point
Tangential plane
Concave surface
Receiver
S_0 (source)

Figure 9-8. Image construction from a concave plane Adapted from Reference 1.

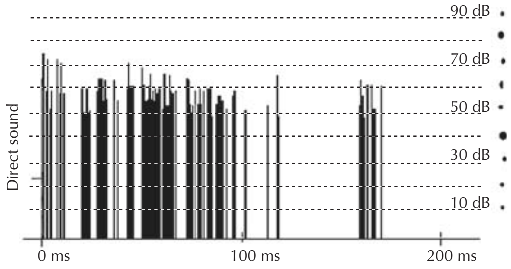


Figure 9-9. A reflectogram (or echogram) display of reflections at a specific point inside a room.

Although it was originally developed solely for the determination of the low-order specular reflections taking place in rectangular rooms due to the geometric increase in the complexity of the computations required, the technique was expanded to predict the directions of the specular reflections from a wide range of shapes.[4] In the image method, the boundaries of the room are effectively being replaced by sources and there is no theoretical limit to the order of reflection that can be handled by the image methodology. From a practical standpoint, the number of the computations that are required tends to grow exponentially with the order of the reflections and the number of surfaces as shown in Eq. 9-10, where N_{IS} represents the number of images, N_W is the number of surfaces that define the room, and i is the order of the reflections.[5]

$$N_{IS} = \frac{N_W}{N_W - 2}[(N_W - 1)^i - 1] \quad (9\text{-}10)$$

Furthermore reflections must be analyzed properly in terms of their visibility to the receiver in the case of complicated room shapes where some elements may block the reflections from the receiver as shown in Fig. 9-10.

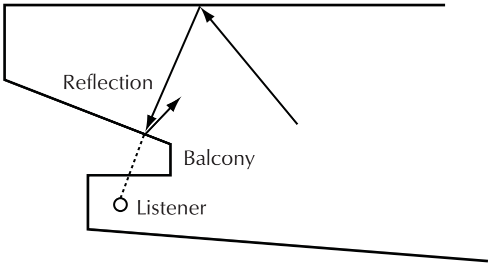


Figure 9-10. The reflection is not visible to the listener due to the balcony obstruction.

The model must also constantly check for the validity of the virtual sources to insure that they actually contribute to the real reflectogram by emulating reflections taking place inside the room and not outside its physical boundaries. Fig. 9-11 illustrates such a situation.

In Fig. 9-11 a real source S_0 creates a first-order image S_1 across boundary 1. This is a valid virtual source that can be used to determine the magnitude and the direction of first-order specular reflections on the boundary surface 1. If one attempts to create a second-order virtual source S_2 from S_1 with respect to boundary surface 2 to find the second order reflection, the image of this virtual source S_2 with respect to boundary 1 is called S_3 but it is contained outside the boundary used to create it and it cannot represent a physical reflection.

Once the map of all the images corresponding to the reflection paths has been stored, the intensity of each individual reflection can be computed by applying Eq. 9-9 introduced earlier. Since the virtual sources do represent the effect of the boundaries on the sound waves, the frequency dependence of the absorption coefficients of the surfaces is modeled by changing the power radiated by the virtual sources; thus, once the image map is obtained the model can be used to rapidly simulate an unlimited number of "what if" simulations pertaining to material changes as long as the locations of the sources and the receiver remain unchanged. A further correction for the air absorption resulting from the wave traveling over extended distances can also be incorporated at this time in the simulation. The same reasoning applies to the frequency distribution of the source: since the image map (and the resulting location of the reflections in the time domain) is a sole function of source and receiver position, the image model can rapidly perform "what if" simulations to yield reflectograms at various frequencies.

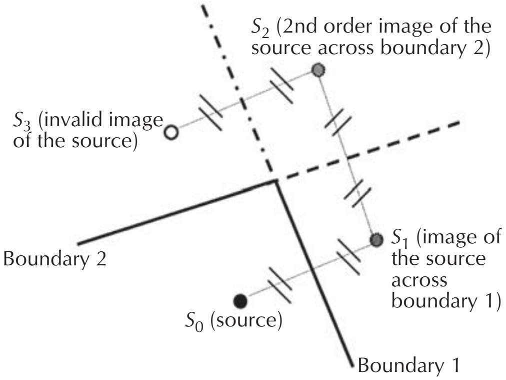

S_2 (2nd order image of the source across boundary 2)
S_3 (invalid image of the source)
Boundary 2
S_1 (image of the source across boundary 1)
S_0 (source)
Boundary 1

Figure 9-11. Invalid images can be created when a virtual source is reflected across the boundary used to create it. Adapted from Reference 4.

The image methodology does not readily account for variations in the absorption coefficient of the surfaces as a function of the angle of incidence of the wave. When taking into account all of the properties in the transmission medium, it can be shown that many materials will exhibit a substantial dependence of their absorption coefficient on the incidence angle of the wave, and in its most basic implementation the image method can misestimate the intensity of the reflections. It is however, possible to incorporate the relationship between angle of incidence and absorption coefficient into a suitable image algorithm in order to yield more accurate results, although at the expense of computational time.

In an image model the user can control the length of the reflection path as well as the number of segments (i.e., the order of the reflections) that comprise it. This allows for a reduction in the computational time of the process since virtual sources located beyond a certain distance from the receiver location can be eliminated while not compromising the fact that all of the reflections within a specific time frame are being recorded, and the image method can lead to very accurate results in the modeling of the arrival time of reflections at a specific location.

Efficient computer implementations of the image methodology have been developed[4] to allow for a fast output of the reflections while also checking for the validity of the images and for the presence of obstacles. Still the method is best suited to the generation of very accurate reflectograms of short durations (500 ms or less) and limited number of reflections (fifth order maximum for typical applications). These factors do not negatively affect the application of the image method in acoustical modeling since in a typical large space—like an auditorium or a theater—the sound field will become substantially diffuse after only a few reflections and some of the most relevant perceived attributes of the acoustics of the space are correlated to information contained in the first 200 ms of the reflectogram.

Ray-Tracing Models. The ray-tracing methodology follows the assumptions of geometrical acoustics presented at the onset of this section, but in this instance the source is modeled to emit a finite number of rays representing the sound waves in either an omnidirectional pattern for the most general case of a point source, or in a specific pattern if the directivity of the source is known. Fig. 9-12 shows an example of a source *S* generating rays inside a space and how some of the rays are reflected and reach the receiver location *R*.

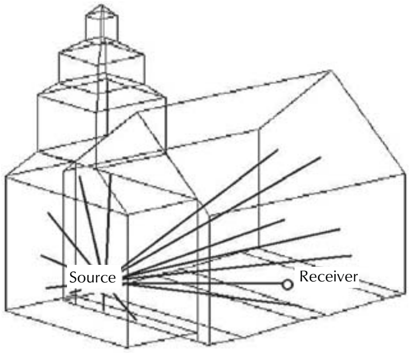


Figure 9-12. Rays are generated by a source, *S*. Some of the rays reach the receiver, *R*.

In this instance, the goal is not to compute all of the reflection paths reaching the receiver within a given time frame but to yield a high probability that a specified density of reflections will reach the receiver (or detector usually modeled as a sphere with a diameter selected by the user) over a specific time window.

In the case of the image method, the boundaries of the room are replaced by virtual sources that dictate the angle of the reflections of the sound waves. In a similar fashion, the ray-tracing technique creates a virtual environment in which the rays emitted by the source can be viewed as traveling in straight paths across virtual rooms until they reach a virtual listener as presented in Fig. 9-13.

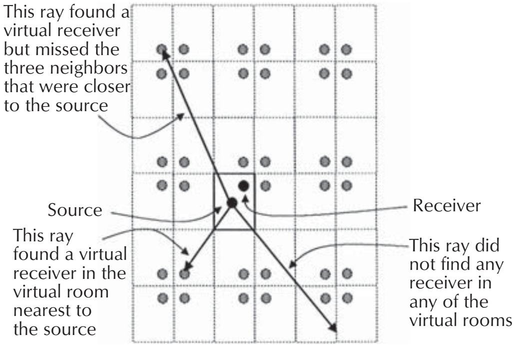

This ray found a virtual receiver but missed the three neighbors that were closer to the source
Source
This ray found a virtual receiver in the virtual room nearest to the source
Receiver
This ray did not find any receiver in any of the virtual rooms

Figure 9-13. The rays can be seen as traveling in straight paths across virtual images of the room until they intercept a receiver. Adapted from Reference 4.

The time of travel and location of the ray are then recorded and can yield a reflectogram similar to that presented earlier in Fig. 9-9.

The main advantage of the ray-tracing technique is that since the model is not trying to find all of the reflection paths between source and receiver, the computational time is greatly reduced when compared to an image technique; for a standard ray-tracing algorithm, the computational time is found to be proportional to the number of rays and to the desired order of the reflections. Another advantage inherent to the technique is that multiple receiver locations may be investigated simultaneously since the source is emitting energy in all directions and the model is returning the number and the directions of rays that are being detected without trying to complete a specific path between source and receiver. On the other hand, since the source is emitting energy in many directions and one cannot dictate what the frequency content of a specific ray is versus that of another, the simulations pertaining to the assessment of frequency-dependent absorption must be performed independently and in their entirety for each frequency of interest.

One problem associated with the ray-tracing technique is that the accuracy of the detection is strongly influenced by size of the detector. A large spherical detector will record a larger number of hits from the rays than another spherical detector of smaller diameter, even if the respective centers of the spheres are located at the exact same point in space. Furthermore, the ray-tracing method may lead to an underestimation of the energy reaching the detector (even if its size is considered adequate) unless large numbers of rays are used since the energy is sampled via rays that diverge as they spread from the source thus increasing the possibility that low-order reflections may miss the detector.

Techniques combining the image methodology and the ray-tracing approach have been developed.[5] The algorithms aim at reducing the number of images to be considered by using the more computationally efficient ray-tracing technique to conduct the visibility test required by the image method.

Beam-Tracing Models. The triangular area that is defined between two adjacent rays emanating from a source is called a 2D ray; more than two rays can also be used to define a 3D pyramidal or conical region of space in which the acoustical energy is traveling away from the source. In these instances, the source is viewed at emitting beams of energy, and the associated modeling techniques are known as *beam-tracing* methods. Figure 9-14 shows an example of a beam and its reflection path from a surface.

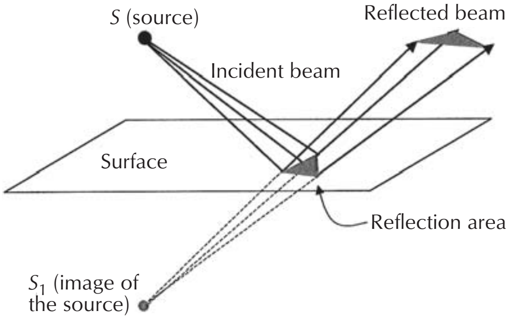


Figure 9-14. A 3D beam is emitted by a source *S* and reflects at a surface.

The beam-tracing technique offers the advantage of guaranteeing that the entire space defining the model will receive energy since the directions of propagations are not sampled as in the case of the traditional ray-tracing approach. Virtual source techniques are used to locate the points that define the reflection zones across the boundaries of the room. On the other hand, the technique requires very complex computations to determine the reflection patterns from the surfaces since the reflection cannot be viewed as a single point as in the case of the ray-tracing technique: when 2D beams are used, the reflections from the surfaces must be considered as lines, while 3D beams define their reflections as areas. Care must also be taken to account for overlapping of the beams by each other or truncation of the beams by obstacles in the room.

Although the computational complexity of the model is substantially increased when it comes to assessing the direction of the reflections, the departure from the single point reflection model presents numerous advantages over the traditional image and/or ray-tracing technique. The issues associated with the divergence of the reflections as a function of increased distance from the source are naturally handled by the beam-tracing approach. Furthermore, the effects of acoustical diffusion can be modeled—at least in an estimated fashion—since the energy contained in the beams can be defined as having a certain distribution over either the length of the intersecting lines (for 2D beams) or areas (for 3D beams). For example, an adaptive beam-tracing model[6] that controls the cross-sectional shape of the reflecting beam as a function of the shape of the reflecting surfaces also allows for an evaluation of the diffuse and specular energy contained inside a reflecting beam. If the energy contained inside the incident beam is E_B and the energy reflected from a surface is E_R, then one can write

$$E_R = E_B(1-\alpha)(1-\delta) \qquad (9\text{-}11)$$

where,

α is the surface's absorption coefficient,

δ is the surface's diffusion coefficient.

The energy E_D that is diffused by the surface is found to be proportional to the area of illumination A and inversely proportional to the square of an equivalent distance L between the source and the reflection area

$$E_D \propto \frac{E_B A \delta(1-\alpha)}{4\pi L^2} \qquad (9\text{-}12)$$

The adaptive algorithm allows for a separate assessment of the specular and of the diffuse reflections from the same geometrical data set that represents the travel map of the beams inside the space. In this instance the diffused energy from a given surface is redirected to other surfaces in a recursive fashion via radiant exchange, a technique also used in light rendering applications. The diffuse and the specular portions of the response can then be recombined to yield a reflectogram that presents a high degree of accuracy, especially when compared to traditional ray-tracing techniques. Fig. 9-15 shows a comparative set of impulse response reflectograms obtained by the adaptive beam tracing, the image, the ray-tracing, and the nonadaptive beam-tracing techniques in a model of a simple performance space containing a stage and a balcony.

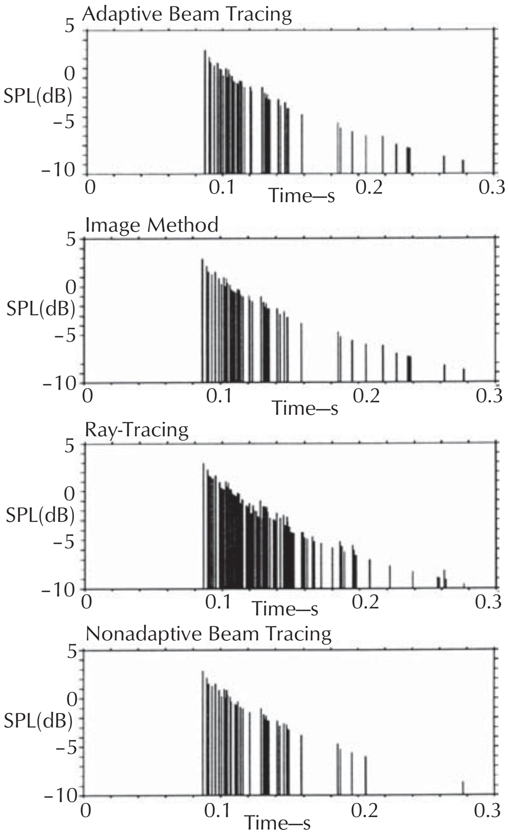


Figure 9-15. Comparative reflectograms for a simple room model. From Reference 6.

The adaptive model is able to yield a reflectogram that is extremely close to that obtained with an image method—i.e., it is able to generate all of the possible reflections paths at a single point in space. From the perspective of computing efficiency, the adaptive-beam tracing methodology compares favorably with the image methodology especially when the complexity of the room and/or the order of the reflections is increased.

Other variants of the beam-tracing approach have been developed. In the *priority-driven* approach,[7] the algorithms are optimized to generate a series of the most relevant beams from a psychoacoustics perspective so that the reflectogram can be generated very rapidly, ideally in real time and the model can be used in an interactive fashion. The beams are ranked in order of importance based on a priority function that aims at accurately reproducing the early portion of the reflectogram since it is by far the most relevant to the perception of the space from the perspective of psychoacoustics. The late portion of the reflectogram (the late reverberation) is then modeled by using a smooth and dense energy profile that emulates the statistical decay of the energy in a large space.

A note on diffusion: The issue of diffusion has been of prime interest to the developers of computer models since commercially available or custom-made diffusers are often integrated into room designs, sometimes at a high cost. Although diffusion is an essential qualitative part of the definition of a sound field, the quantitative question of “how much diffusion is needed?” is often answered using considerations that have little foundation in scatter theory and/or general acoustics. A concept as elementary as reverberation finds its classical quantitative representation (the Sabine/Eyring equation and associated variants) rooted into the notion that the sound field is assumed to be diffuse, and unless this condition is met in reality, one will encounter substantial errors in predicting the reverberation time. Today’s advanced large room computer acoustic simulation software products incorporate the ability to model diffused reflections using either a frequency dependence function or an adaptive geometry that spreads out the incident energy of the sound ray over a finite area. This allows for a much more accurate correlation between predicted and test data especially in rooms that have geometry involving shapes and aspect ratios that are out of the ordinary, noneven distribution of absorptive surfaces, and/or coupled volumes.[8] Under these conditions the incorporation of diffusion parameters into the model is necessary and a specular-only treatment of the reflections (even when using an efficient ray-tracing technique) will lead to errors.

9.2.2.2 Wave Equation Models

Wave equation models are based on an evaluation of the *fundamental wave equation*, which in its simplest form relates the pressure p of a wave at any point in space to its environment via the use of the 3D Laplacian operator ∇^2 and the wave number k:

$$\nabla^2 p + k^2 p = 0 \tag{9-13}$$

Solving the fundamental wave equation allows for an exact definition of the acoustic pressure at any specific point since appropriate boundary conditions defining the physical properties of the environment (surfaces, medium) can be used whenever required. As an example, in a model based on the wave equation the materials that comprise the environment (like the room) can be defined in terms of their *acoustical impedance* z given by

$$\mathbf{z} = \frac{\mathbf{p}}{\mathbf{U}} \tag{9-14}$$

where,

p refers to the pressure of the wave,

U to its velocity in the medium.

When using the wave equation, issues having to do with diffraction, diffusion, and reflections are automatically handled since the phenomena are assessed from a fundamental perspective without using geometrical simplifications. The main difficulty associated with the method is found in the fact that the environment (surfaces and materials) must be described accurately in order for the wave equation to be applied: either an analytical or a numerical approach can be used to achieve this goal.

9.2.2.2.1 Analytical Model: Full-Wave Methodology.

An analytical model aims at providing a mathematical expression that describes a specific phenomenon in an accurate fashion based on underlying principles and/or physical laws that the phenomenon must obey. Because of this requirement the analytical expression governing the behavior of a model must be free of correction terms obtained from experiments and of parameters that cannot be rigorously derived from—or encountered in—other analytical expressions.

The complexity of the issues associated with sound propagation has prevented the development of a single and unified model that can be applied over the entire range of frequencies and surfaces that one may encounter in acoustics; most of the difficulties are found in trying to obtain a complete analytical description of the scattering effects that take place when sound waves impinge on a surface. In the words of J.S. Bradley,[9] one of the seminal researchers in the field of architectural acoustics:

> *Without the inclusion of the effects of diffraction and scattering, it is not possible to accurately predict values of conventional room acoustics parameters [...]. Ideally, approximations to the scattering effects of surfaces, or of diffraction from finite size wall elements should be derived from more complete theoretical analyses. Much work is needed to develop room acoustics models in this area.*

In this section, we present the full-wave methodology,[10] one analytical technique that can be used for the modeling of the behavior of sound waves as they interact with nonidealized surfaces, resulting in some of the energy being scattered, reflected, and/or absorbed as in the case of a real space. Due to the complexity of the mathematical foundation associated with this analytical technique, only the general approach is introduced here and the reader is referred to the bibliography and reference section for more details.

The full-wave approach (originally developed for electromagnetic scattering problems) meets the important condition of acoustic reciprocity requiring that the position of the source and of the receiver can be interchanged without affecting the physical parameters of the environment like transmission and reflection coefficients of the room's surfaces. In other words if the environment remains the same, interchanging the position of a source and of a receiver inside a room will result in the same sound fields being recorded at the receiver positions. The full-wave method also allows for exact boundary conditions to be applied at any point on the surfaces that define the environment, and it accounts for all scattering phenomena in a consistent and unified manner, regardless of the relative size of the wavelength of the sound wave with that of the objects in its path. Thus the surfaces do not have to be defined by general (and less than accurate) coefficients to represent absorption or diffusion, but they can be represented in terms of their inherent physical properties like density, bulk modulus, and internal sound velocity.

The full-wave methodology computes the pressure at every point in the surface upon which the waves are impinging, and follows the shape of the surface. The coupled equations involving pressure and velocity are then converted into a set of equations that separate the forward (in the direction of the wave) and the backward components of the wave from each other, thus allowing for a detailed analysis of the sound field in every direction. Since the full-wave approach uses the fundamental wave equation for the derivation of the sound field, the model can return variables such as sound pressure or sound intensity as needed.

The main difficulty associated with the full-wave method is that the surfaces must also be defined in an analytical fashion. This is possible for simple (i.e., planar, or curved) surfaces for which equations are readily available, but for more complicated surfaces—such as those found in certain shapes of diffusers—an analytical description is more difficult to achieve, and the methodology becomes restricted to receiver locations that are located at a minimum distance from the surfaces upon which the sound waves are impinging. Still for many problems the full-wave methodology is a very accurate and efficient way to model complicated scattering phenomena.

9.2.2.3 Numerical Model: Boundary Element Methodology

The *boundary element analysis* (BEA) techniques are numerical methods that yield a quantitative value of the solution to the problem under investigation. BEA techniques[11,12,13,14] can be used in solving a wide range of problems dealing with the interaction of energy (in various forms) with media such as air and complex physical surfaces, and they are well suited to the investigation of sound propagation in a room. Although the method is based on solving the fundamental differential wave equation presented earlier, the BEA methodology makes use of an equivalent set of much simpler algebraic equations valid over a small part of the geometry, and then expands the solution to the entire geometry by solving the resulting set of algebraic equations simultaneously. In essence, the BEA technique replaces the task of solving one very complex equation over a single complicated surface by that of solving a large quantity of very simple equations over a large quantity of very simple surfaces. In a BEA implementation the surface is described using a meshing approach as shown in Fig. 9-16.

In the BEA method the analytical form of the solution over the small domain (area) of investigation is not directly accessible for modification. The use exercises control over the solution by properly specifying the domain (geometry) of the problem, its class (radiation or scattering), the parameters of the source (power, directivity, location), and, of course, the set of boundary conditions that must be applied at each area defined by the mesh. It is thus possible to assign individual material properties at every location in the mesh of the model in order to handle complex scattering and absorption scenarios, if needed. Although it can be adapted to solving acoustical problems in the time domain the BEA technique is better suited to providing solutions in the frequency domain since the characteristics of the materials are considered to be time-invariant but frequency dependent.

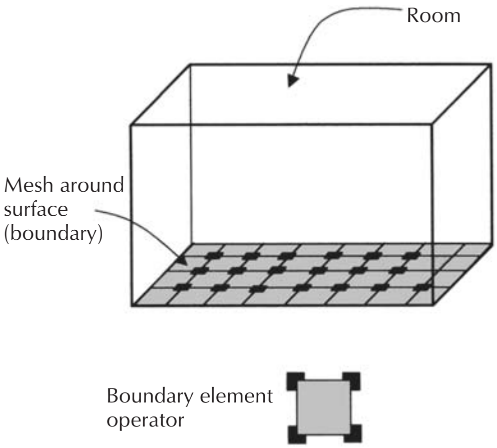


Figure 9-16. A mesh describes a boundary in the BEA method.

The main issue that is associated with the use of BEA methodology for the investigation of acoustical spaces is that the size of the elements comprising the mesh representing the surfaces dictates the accuracy of the solution. A small mesh size will, of course, allow for a very accurate description of the surfaces, both geometrically and in terms of its materials, but it will also drastically affect the computational time required to yield a solution. On the other hand, a large mesh size will yield very fast results that may be inaccurate because the algebraic equations that are used in lieu of the fundamental wave equation improperly being applied over large surfaces. A comparison of the accuracy yielded by BEA techniques over very simple geometries indicates that a minimum ratio of seven to one (7:1) must exist between the wavelength of the sound and the element size in order to bind the dependence of the BEA analysis on the size of its element to less than a ±0.5 dB resolution. In other words, the wavelengths considered for analysis must be at least seven times larger than the largest mesh element in order for the methodology to be accurate. For this reason the BEA methodology is very efficient and accurate to model sound propagation at low frequencies (below 1000 Hz), but it becomes tedious and cumbersome at higher frequencies since in this instance the mesh must be modeled with better than a 50 mm resolution. Still the technique can be shown to yield excellent results when correlating modeled projection and actual test data from complicated surfaces such as diffusers.[13]

Numeric Model: Finite Difference Time-Domain Methodology. As mentioned earlier, the BEA techniques are best suited to yielding data in the frequency domain, although they can be adapted to provide time-domain information albeit at a cost in computing efficiency. Another numerical methodology that uses a discrete representation of the acoustical environment is known as *Finite-Difference Time-Domain* (FDTD) and it is very efficient in terms of computational speed and storage while also offering excellent resolution in the time domain. It has been demonstrated[15] that the technique can be used to effectively model low-frequency problems in room acoustics simulations and the results are suitable for the generation of reflectograms.

In the finite difference (FD) approach instead of describing the surface with a mesh (as with the BEA technique), a grid is used and the algebraic equations are solved at the points of the grid as shown in Fig. 9-17.

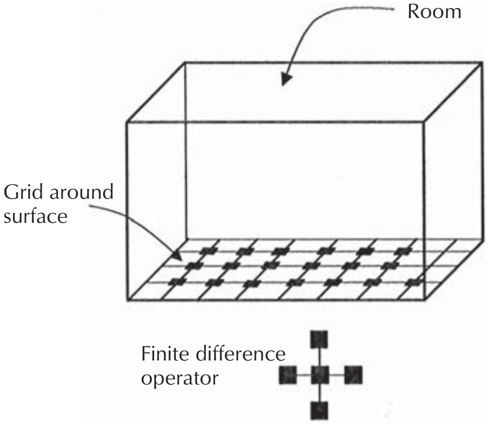


Figure 9-17. A grid describes a surface in the FD method.

In this instance the size of the grid can be made as small as needed to provide a high degree of resolution when needed and the grid points can be defined using the most effective coordinate system for the application. For example, a flat surface could be defined with a (x, y, z) Cartesian grid system while cylinders (for pillars) and spheres (for audience's heads) could be expressed with cylindrical and spherical systems respectively.

9.2.2.4 Statistical Models

The use of statistics in acoustical modeling is primarily reserved for the study of the behavior of sound in rectangular and rigid rooms where the dominant phenomena that are taking place are related to modes. The issues of modal frequencies, modal density, and mode distributions are presented, along with the appropriate descriptive equations in Chapter 5—Small Room Acoustics.

Another application of statistics in acoustical modeling can be found in situations where resonance effects take place at high frequencies, as opposed to the traditionally low frequencies associated with room modes. A technique known as Statistical Energy Analysis[16] (SEA) can be used to accurately account for the effect of modal resonance effects that take place in systems such as partitions and walls, by analyzing the kinetic energy and the strain energy associated with vibrating structures. An SEA model will describe a vibrating system (such as a wall) with mass and spring equivalents and will allow for the analysis of the effect that adding damping materials will have on the vibration spectrum. SEA techniques are optimized for frequency-domain analysis and the output cannot be used for time-domain applications to add information to the impulse response of a room, or to yield a reflectogram; still, the main advantage of SEA is that real materials such as composite partitions with different degrees of stiffness, construction beams, and acoustical sprays can be modeled in a precise manner (i.e., not only in terms of a unique physical coefficient) over an extended range of frequency.

9.2.2.5 Small Room Models

A room that is acoustically small can be defined as one in which classically defined reverberation phenomena (using the assumption of a diffuse sound field) do not take place, but rather that the sound decays in a nonuniform manner that is a function of where the measurement is taken. The use of diffusion algorithms in large room models that rely on either ray-tracing, image source, or adaptive algorithms has vastly improve the reliability of the prediction models in a wide range of spaces, however accurate predictions of the sound field can be made in small rooms considering the interference patterns that result from modal effects. Figs. 9-18 and 9-19 shows the mapping[17] of the interference patterns resulting from modal effects in an 8 m × 6 m room where two loudspeakers are located at B1 and B2. In the first instance, a modal effect at 34.3 Hz creates a large dip in the response at about 5.5 m, while the second case shows a very different pattern at 54.4 Hz. Such models are very useful to determine the placement of low-frequency absorbers into a room in order to minimize the impact of modal effects at a specific listening location, and they are a good complement to the large room models that typically do not investigate the distribution of the sound field at the very low frequencies.

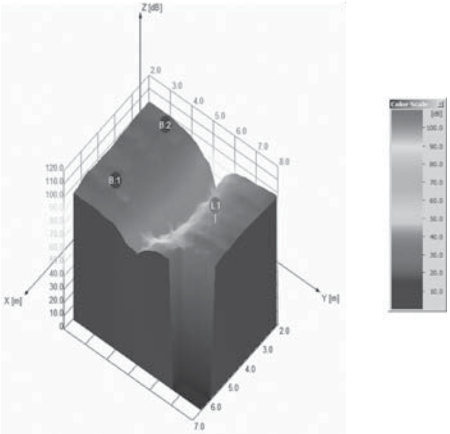


Figure 9-18. A mapping of the sound field resulting from modal interference patterns at 34.3 Hz. From CARA.

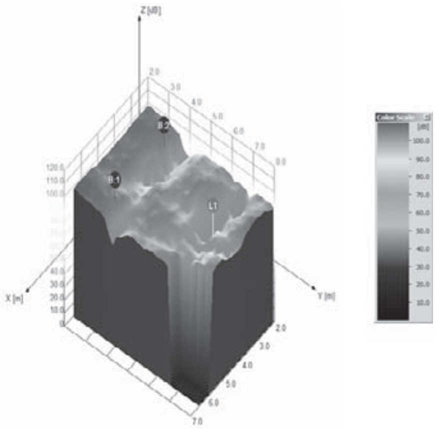


Figure 9-19. A mapping of the sound field resulting from modal interference patterns at 54.4 Hz. From CARA.

9.2.3 Empirical Models

Empirical models are derived from experiments and are described by equations that typically follow curve-fitting procedures of the data obtained in the observations. No analytical, geometrical, and/or statistical expression is developed to fully explain the interdependence of variables and parameters in the model, but a general form of a descriptive expression may be constructed from underlying theories. Empirical models have been extensively used for many years in acoustical modeling due to the large quantity of variables and parameters that are often present when dealing with issues of sound propagation in a complicated environment, and this section will present only a couple of examples.

9.2.3.1 Gypsum Cavity Wall Absorption

There is numerous test data available pertaining to the sound transmission class (STC) of various wall constructions, but little has been investigated regarding the sound absorption of walls constructed of gypsum (drywall) panels. The absorption of a composite wall panel is partly diaphragmatic (due to the mounting), partly adiabatic (due to the porosity of the material and the air), and some energy is lost inside the cavity via resonance. The complicated absorption behavior of gypsum walls has been described[18] using an empirical model that takes into account absorption data acquired in reverberation chamber experiments. The mathematical model is fitted to the measured data to account for the resonant absorption of the cavity by assuming that the mechanical behavior of the wall can be modeled by a simple mechanical system.

In this model, the resonance frequency at which the maximum cavity absorption takes place is given by

$$f_{MAM} = P\sqrt{\frac{m_1 + m_2}{d(m_1 m_2)}} \qquad (9\text{-}15)$$

where,

m_1 and m_2 are the mass of the gypsum panels comprising the sides of the wall in kg/m^2,

d is the width of the cavity expressed in millimeters,

P is a constant with the following values:

If the cavity is empty (air), *P* = 1900,

If the cavity contains porous or fibrous sound-absorptive materials, *P* = 1362.

The empirical model combines the maximum absorption α_{MAM} taking place at the resonant frequency given by Eq. 9-15 with the high-frequency absorption α_S into a form that fits data obtained experimentally, to give an equation that allows for the prediction of the absorption coefficient of the wall as a function of frequency:

$$\alpha(f) = \alpha_{MAM}\left(\frac{f_{MAM}}{f}\right)^2 + \alpha_{MAM} \qquad (9\text{-}16)$$

Although it does not take into account all of the construction variables (stud spacing, bonding between layers) the model still provides accurate prediction of the sound absorption parameters of various gypsum wall constructions.

9.2.3.2 Absorption from Trees and Shrubs

When dealing with issues related to outdoor noise propagation one may need to predict the anticipated noise reduction that can be expected from vegetation. In this instance, some of the general attributes of a tree-barrier such as height and width can be modeled from geometry, but others like leaf density, wind resistance, or diffraction effects from trunks may prove very difficult to describe either analytically or geometrically. In this instance an empirical model that fits experimental data to polynomial equations based on statistical regression is the most appropriate[19] to yield the sound pressure level at various distances from a sound source while taking into account tree height, width of the tree barrier, wind velocity, and tree type. An example of such an equation is presented below, and it is shown to give excellent (±1 dB) accuracy between predicted and observed levels for distances extending 150 ft to 400 ft from the source that is assumed to be truck noise on an interstate highway. The receiver is shielded from the traffic by a belt of conifer trees planted along the interstate.

$$L_{dB} = 81.65 - 0.2257H - 0.0229W + 0.728V - 0.0576D \tag{9-17}$$

where,

L_{dB} is the predicted sound level behind the tree belt,

H is the height of the tree belt, expressed in feet,

W is the width of the tree belt, expressed in feet,

V is the wind velocity component in the direction of the sound propagation, expressed in mph,

D is the distance from the receiver to the tree belt.

Other equations are available for different sources and different types of trees. In this class of empirical models, no attempt is made to support the equation by analytical expressions but this does not affect the usefulness or the accuracy of the model.

9.2.4 Hybrid Models

As the name implies *hybrid models* use a combination of techniques to yield results and the choice of the technique may be based on a specific need such as fast output, accuracy, range of applicability, etc. A hybrid model can combine the inherent accuracy of the image method for the determination of reflection arrival time in the specular case, with an adaptive beam-tracing approach when diffusion is required, and may also incorporate some BEM computations for complicated materials wherever required. A hybrid model can also rely on empirical approaches to provide a confidence factor for results obtained from physical scale models or from statistical approaches.

An example of hybrid techniques can be found in models that are aimed at assessing outdoor noise propagation.[20] In this instance, the objects that are in the path of the sound waves are typically buildings or large natural obstacles and can be considered to be much larger than the wavelength, except for the lowest frequencies, and as such, the geometrical acoustics assumptions apply very well; as such, the image method is very appropriate to compute reflection paths between obstacles. On the other hand one cannot ignore the fact that outdoor noise may contain a lot of low frequencies and that diffraction effects will take place; in this instance the model must use an appropriate description of diffraction such as the one presented in Chapter 4—Acoustical Treatment of Rooms and the model may also be refined from empirical data table to represent complicated sources such as car traffic, aircraft, and trains since point source assumptions become invalid and the sources are also moving. Figs. 9-20 and 9-21 shows the type of noise prediction maps that can be obtained from such a model; in the first instance the noise sources are a combination of street traffic and large mechanical systems, and the model takes into account the diffraction effects of various buildings. In the second instance, the model is used to assess the difference in expected noise levels between different types of pavements (asphalt vs. concrete) based on traffic data on a segment of road that is surrounded by residences.

Hybrid models are also found in construction applications, where they combine analytical techniques based on specific equations with databases of test results obtained in the laboratory and in the field. As an example, a simple model could be developed using the Mass Law in order to predict the sound transmission between two spaces and yield an estimate of the Sound Transmission Class (STC) of the partition, however, the results would not be very useful because they would extensively be influenced by the construction technique and the presence of flanking paths. With a model that takes into account the construction techniques of the partition,[21] the results are much more accurate and provide the designer with valuable insight on the weak links of the construction as they pertains to noise transmission.

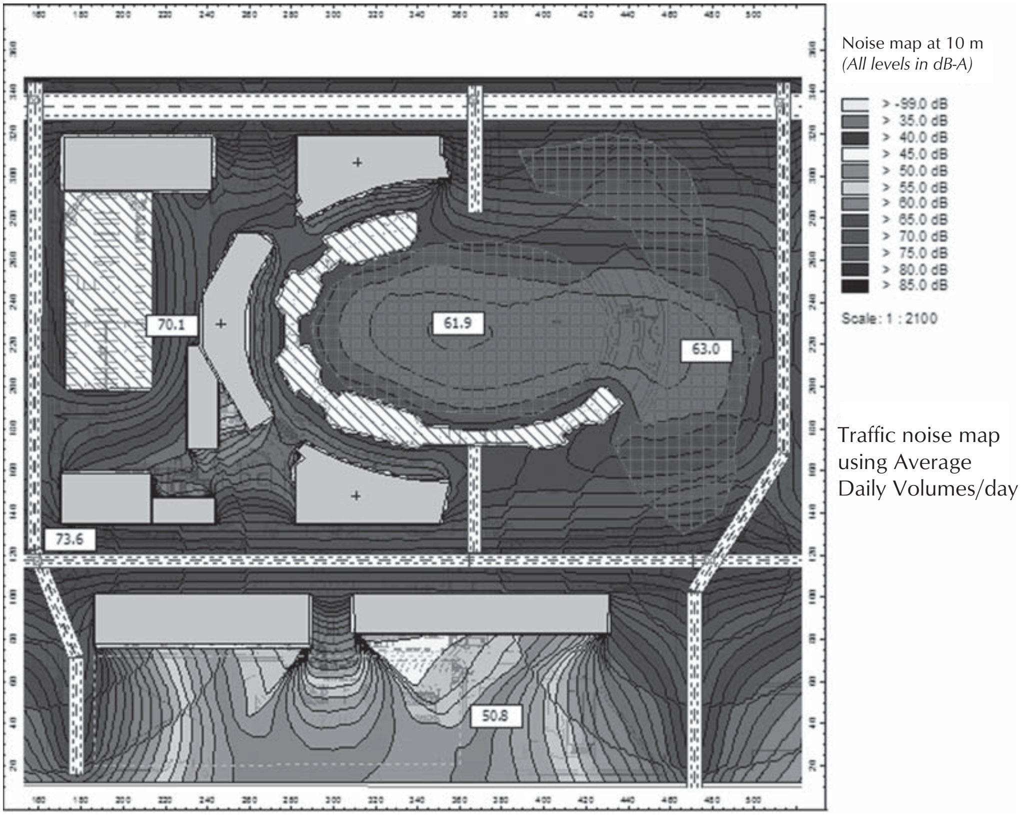

Noise map at 10 m
(All levels in dB-A)
> -99.0 dB
> 35.0 dB
> 40.0 dB
> 45.0 dB
> 50.0 dB
> 55.0 dB
> 60.0 dB
> 65.0 dB
> 70.0 dB
> 75.0 dB
> 80.0 dB
> 85.0 dB
Scale: 1 : 2100
70.1
61.9
63.0
73.6
50.8
Traffic noise map
using Average
Daily Volumes/day

Figure 9-20. A noise map from an outdoor propagation model. CadnaA by DataKustik GmbH.

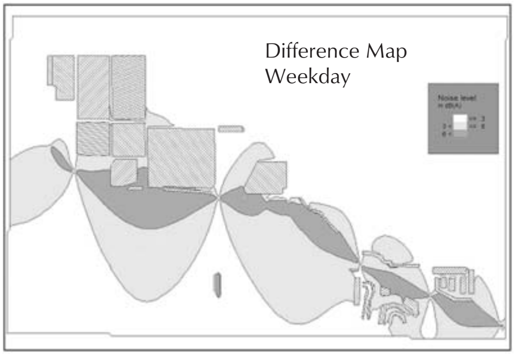

Difference Map
Weekday
Noise level
in dB(A)

Figure 9-21. Difference maps for noise generated by different types of road pavements. Courtesy SoundPLAN, by SoundPLAN LLC.

9.3 Auralization

> *Auralization is the process of rendering audible, by physical or mathematical modeling, the sound field of a source in a space, in such a way as to simulate the binaural listening experience at a given position in the modeled space.*[22]

Auralization systems have been in existence since the 1950s.[23] During the early experiments, researchers used a physical 1:10 scale model in which a tape containing speech and music samples was played back at scaled-up speed through a scaled omnidirectional source while also taking into account the air absorption and scaling the reverberation time of the model. A recording of the sound was made at the desired receiver locations using a scaled dummy head and the results were played back at a scaled-down speed under anechoic conditions using two speakers. The sound of the model was then subjectively assessed and compared to that perceived in the real room.

The technique—or variants of it—was used for the prediction of the acoustics of both large and small rooms throughout the 1970s, however, with computer systems becoming increasingly faster and more affordable auralization techniques based on computational models have been developed to yield an audible representation of the sound field at any specified receiver location by using the results of the acoustical modeling phase. Various implementations of auralization have been put into place[24,25,26,27] at the time of this writing but because of the tremendous developments that are taking place in the field of auralization this section will only explore the general concepts associated with the topic of auralization since it is safe to say that specific imple- mentations of auralization techniques will be subject to changes and additions dictated by new technological advances and/or by market demands.

9.3.1 The Basic Auralization Process

The basic auralization process associated with an acoustical model is illustrated in Fig. 9-22.

The process starts with the reflectograms representing the *impulse response* (IR) of the model obtained at a specific receiver location for various frequencies. The reflectograms contain the information pertaining to the intensity and the direction of arrival of the reflections over a period of time that is deemed suitable to record the desired order and length of the reflections, and they are obtained from any of the methodologies presented in the modeling portion of this chapter. The reflectograms are then convolved—or mixed—with a dry (anechoic) recording of speech or music that can be played back under controlled conditions, using either headphones or loudspeakers, for the purpose of subjective evaluation.

9.3.2 Implementation

The energy reaching the listener is comprised of the direct sound, of the early reflections, and of the late reflections as shown in Fig. 9-23.

The direct sound is easily found and modeled accurately in the reflectogram since it represents the energy traveling from source to receiver in a direct line of sight. The only concern for the accurate auralization of the direct sound is to insure that the attenuation follows the inverse-distance spreading law as dictated by the source configuration and directivity. The early reflections are also obtained from the modeling phase but the reflectogram must be limited in length—or in the order of the reflections—because of computational constraints. The late portion of the reflectogram is usually modeled from a dense and random pattern of reflections with a smooth decay and a frequency content patterned after the reverberation time of the room estimated at various frequencies.

Since the reflectogram typically represents the impulse response at a single point (or within a small volume) in the modeled space, it must be modified in order to represent the binaural sound that would be reaching the eardrums of a listener by and at this point, two separate approaches are available.

9.3.2.1 Binaural Reproduction Using Loudspeakers

The impulse response is divided into its left and right components corresponding to the vertical left and right planes crossing the receiver location, and thus yielding the *binaural impulse response* (BIR) of the room for a listener at the receiver location. The anechoic signal is convolved separately for the left and the right channel, and the result is presented under anechoic and/or near field conditions to a listener using loudspeakers as shown in Fig. 9-24.

This technique has the advantage of being efficient from a computational standpoint since the process is limited to the separation of the IR into the BIR and the resulting convolution into left and right channels for the playback system. The drawback of the technique is that the playback requires a controlled environment where the listener has to maintain a fixed position with respect to the playback system and the crosstalk between the loudspeakers must be very small in order to yield the proper sense of spatial impression.

9.3.2.2 Binaural Reproduction Using Headphones

In this approach, the BIR is further modified by the application of *head-related transfer functions* (HRTF) that represent the effects that the head, torso, shoulders, and ears will have on the sound that reaches the eardrums of the listener. It has been shown[28,29] that these parameters have a drastic influence on the localization of the sound and on its overall subjective assessment. As shown in Fig. 9-25, the reproduction system must now use headphones since the effects of the body and head shape of the listener have already been taken into account. The advantage of this approach is that the playback system is very simple; good quality headphones are readily available and no special setup is required. The drawback is that the implementation of the modified BIR takes time due to the computational requirements for the application of the HRTF. It must also be noted that the HRTF may not accurately describe the specific parameters that a given listener experiences, although current HRTF research has yielded accurate

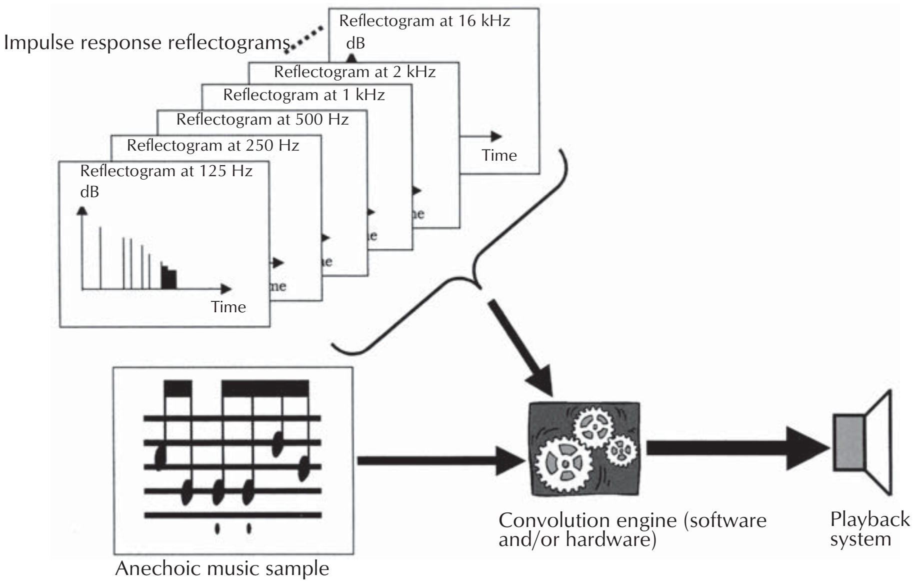

Impulse response reflectograms
Reflectogram at 16 kHz
dB
Time
Reflectogram at 2 kHz
Reflectogram at 1 kHz
Reflectogram at 500 Hz
Reflectogram at 250 Hz
Reflectogram at 125 Hz
dB
Time
Convolution engine (software and/or hardware)
Playback system
Anechoic music sample

Figure 9-22. The basic auralization process.

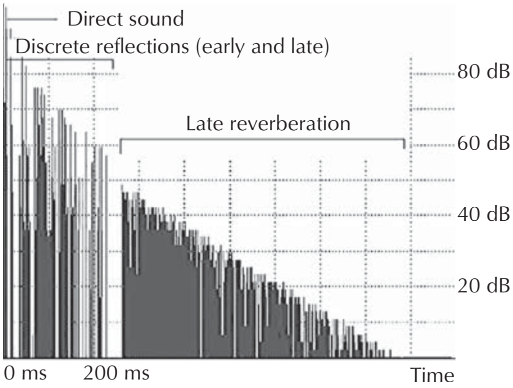


Figure 9-23. An example of a complete reflectogram at 1000 Hz.

composite data for a wide segment of the test results. Another issue of concern when using headphone reproduction is that the apparent source location will move with the listener's head movements, something that does not take place in the real world.

9.3.2.3 Multichannel Reproduction Using Loudspeakers

In this instance, the impulse response of the room is divided into components that correspond to the general locations in space from where the reflections originate as shown in Fig. 9-26. Various systems have been developed[30,31] throughout the years ranging from just a few speakers to hundreds of units driven by dozens of separate channels. The advantage of the technique is that the system relies on the listener's own HRTF while also allowing for head tracking effects. From the perspective of efficiency the approach can be implemented with minimal hardware and software since the reflections can be categorized in terms of their direction of arrival while the IR is being generated. The multichannel reproduction technique can actually be implemented from a physical scale model without the need for computer tools by using delay lines and an analog matrix system.[1] The reproduction system is, of course, rather complicated since it requires substantial hardware and an anechoic environment.

9.3.3 Real-Time Auralization and Virtual Reality

A real-time auralization system allows the user to actually move within the room and to hear the resulting changes in the sound as they actually happen. This approach requires the near-instantaneous computation of the impulse response so that all parameters pertaining to the direct sound and to the reflections can be computed. In a recent implementation[32] the space is modeled using an enhanced image method approach in which a fast ray-tracing preprocessing step is taken to check the visibility of the reflections at the receiver location. The air absorption and the properties of the surface materials are modeled using efficient digital filters, and the late reverberation is described using techniques that give a smooth and dense reflection pattern

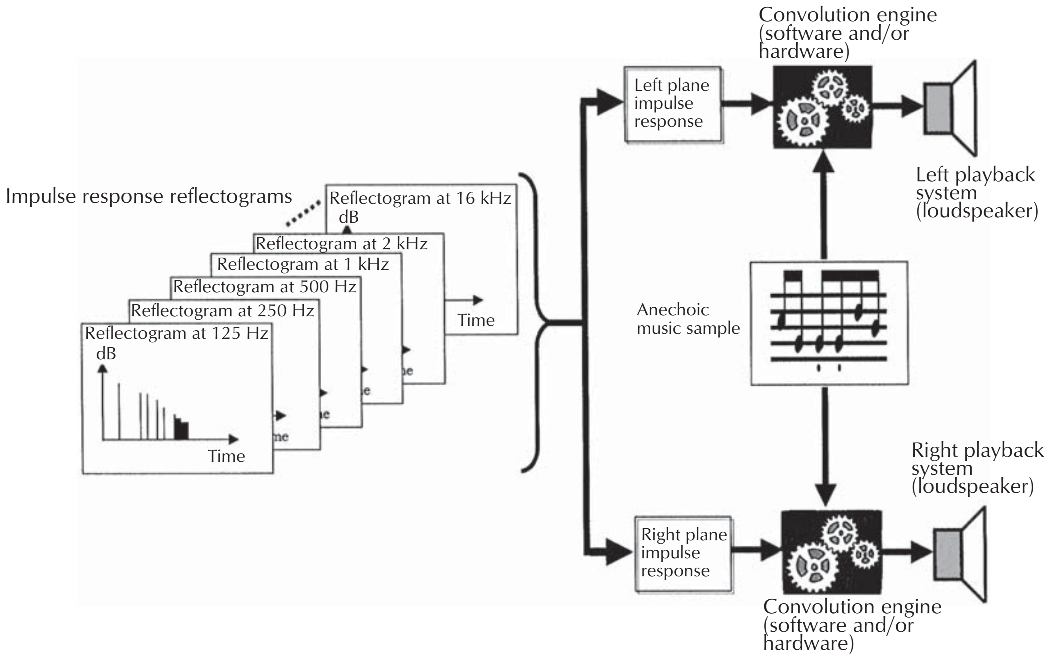

Convolution engine (software and/or hardware)
Left plane impulse response
Left playback system (loudspeaker)
Impulse response reflectograms
Reflectogram at 16 kHz
dB
Reflectogram at 2 kHz
Reflectogram at 1 kHz
Reflectogram at 500 Hz
Reflectogram at 250 Hz
Reflectogram at 125 Hz
dB
Time
Time
Anechoic music sample
Right playback system (loudspeaker)
Right plane impulse response
Convolution engine (software and/or hardware)

Figure 9-24. Auralization with a binaural impulse response and speaker presentation.

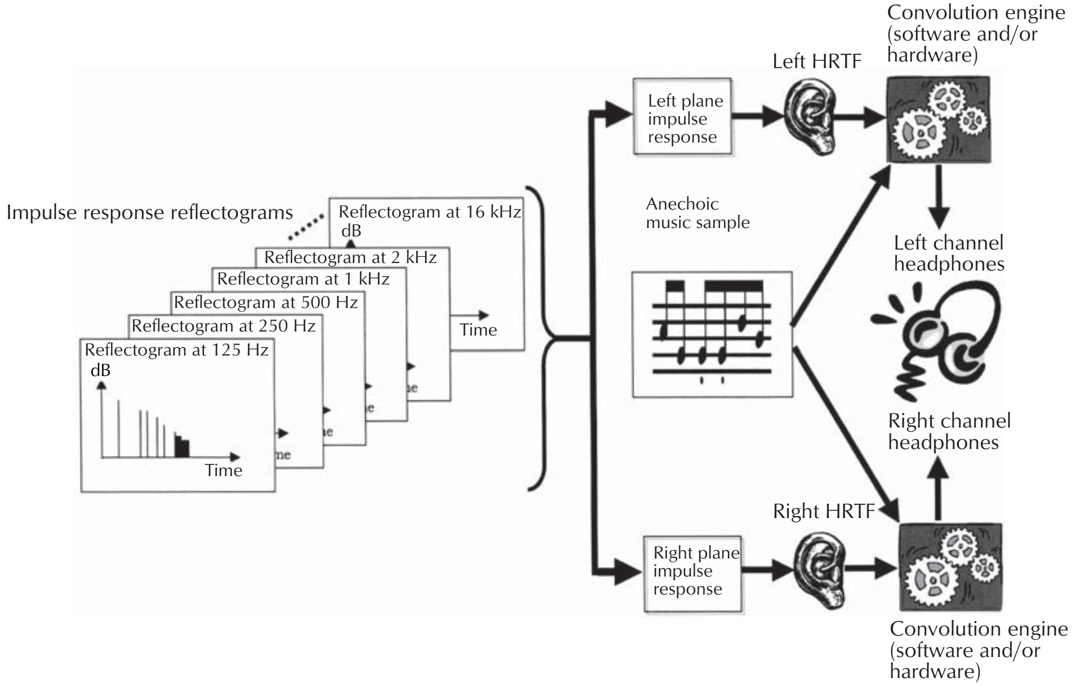

Convolution engine (software and/or hardware)
Left HRTF
Left plane impulse response
Impulse response reflectograms
Reflectogram at 16 kHz
dB
Reflectogram at 2 kHz
Reflectogram at 1 kHz
Reflectogram at 500 Hz
Reflectogram at 250 Hz
Reflectogram at 125 Hz
dB
Time
Time
Anechoic music sample
Left channel headphones
Right channel headphones
Right HRTF
Right plane impulse response
Convolution engine (software and/or hardware)

Figure 9-25. Auralization with a HRTF binaural impulse response and headphone presentation.

that follows the statistical behavior of sound in a bounded space. The technique yields a *parametric room impulse response* (PRIR) in which a combination of real-time and nonreal-time processes performs a model-

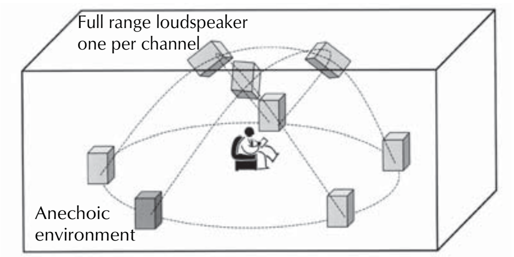


Figure 9-26. Speaker arrangement for multichannel presentation. Adapted from Reference 29.

ing of the physical parameters that define the space. A diagram of the modeling and auralization process of this system is presented in Fig. 9-27.

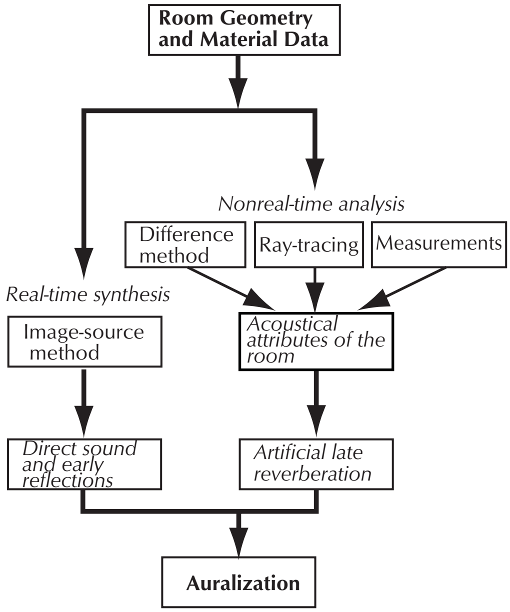

Room Geometry and Material Data
Nonreal-time analysis
Difference method
Ray-tracing
Measurements
Real-time synthesis
Image-source method
Acoustical attributes of the room
Direct sound and early reflections
Artificial late reverberation
Auralization

Figure 9-27. A real-time interactive modeling and auralization system. Adapted from Reference 32.

In this approach, known as *dynamic auralization*, the presentation of the sound field can be done either via binaural headphones or by multi-channel speaker techniques and the auralization parameters must be updated at a fast rate (typically more than ten times per second) in order for the rendering to be of high quality. The impulse response that is used for the convolutions can be a combination of an accurate set of binaural responses (that map head-tracking movements) to account for the early portion of the reflections with a simpler static impulse response that provides the foundation for the calculation of the late part of the sound field. This approach is moderately efficient in terms of computational time and memory consumption and recent developments[33] have been aimed at making use of an efficient means to process the impulse response of a space. Using an approach known as Ambisonics B-format[34] the sound information is encoded into four separate channels labeled W, X, Y and Z. The W channel would be equivalent to the mono output from an omnidirectional microphone while the X, Y and Z channels are the directional components of the sound in front-back (X), left-right (Y), and up-down (Z) directions. This allows a single B-format file to be stored for each location to account for all head motions at this specific location and to produce a realistic and fast auralization as the user can move from one receiver location to the other and experience a near-seamless simulation even while turning his/her head in the virtual model.

9.4 Conclusion

Acoustical modeling and auralization are topics of ongoing research and development. Originally planned for the evaluation of large rooms, the techniques have also been used in small spaces[35] and in outdoor noise propagation studies[36] and one can expect to witness the standard use of these representation tools in a wide range of applications aimed at assessing complex acoustical quantifiers. Even simple digital processing systems such as those offered as plug-ins for audio workstations can be used to illustrate the effect of frequency-dependent transmission loss from various materials using simple equalization and level settings corresponding to octave or third-octave band reduction data.

Further work is needed in the representation and modeling of complicated sources such as musical instruments, automobiles, trains, and other forms of transportation; work is also ongoing in the definition of materials and surfaces so that the effect of vibrations and stiffness is accounted for. Still, the models are rapidly becoming both very accurate and very efficient and they are demonstrating their adequacy at illustrating the complicated issues that are associated with sound propagation and, eventually, sound perception.

References

1. Makrinenko, Leonid I. *"Acoustics of Auditoriums in Public Buildings."* First Publication: 1986 (Russian). Acoustical Society of America (1994) Translated by R.S. Ratner. J.S. Bradley, *Editor.*
2. Pierce, Allan D. *Acoustics. "An Introduction to Its Physical Principles and Applications."* Acoustical Society of America, 1991, 2nd printing.
3. Allen, J. B. and Berkley, D. A. "Image Method for Efficiently Simulating Small-Room Acoustics." *Journal of the Acoustical Society of America,* 65 (4). 943-950, 1979.
4. Barish, J. "Extension of the Image Method to Arbitrary Polyhedra." *Journal of the Acoustical Society of America,* 75 (6). 1827-1836, 1984.
5. Vorländer, M. "Simulation of the Transient and Steady-State Sound Propagation in Rooms Using a New Combined Ray-Tracing/Image Source Algorithm." *Journal of the Acoustical Society of America,* 86 (1). 172-178, 1989.
6. Drumm, I. A., and Lam, Y. M. "The Adaptive Beam Tracing Algorithm."*Journal of the Acoustical Society of America,* 107 (3). 1405-1412, 2000.
7. Min, P., and Funkhouser, T. "Priority-Driven Acoustic Modeling for Virtual Environments." *Eurographics* 2000, 19 (3), 2000.
8. Dalenbäck Bengt-Inge. "The rediscovery of diffuse reflection in room acoustics prediction." A presentation to the 144th Meeting of the Acoustical Society of America, Cancun, 2002.
9. Bradley, J. S. *"Architectural research today."*
10. Bahar, E. "Acoustic Scattering by Two-Dimensionally Rough Interfaces between Dissipative Acoustic Media. Full-Wave, Physical Acoustics, and Perturbation Solutions." *Journal of the Acoustical Society of America,* 89 (1). 19-26, 1991.
11. Kane, J. *Boundary Elements Analysis,* Prentice Hall, 1994.
12. COMET/Acoustics, *User Document V.2.1.* Automated Analysis Corporation, 1994.
13. Cox, T. J. "Predicting the Scattering from Diffusers Using 2-D Boundary Element Methods." *Journal of the Acoustical Society of America,* 96 (1). 874-878, 1994.
14. D'Antonio, P., and Cox, T. "Two Decades of Sound Diffusor Design and Development. Part 2: Prediction, Measurements, and Characterization." *Journal of the Audio Engineering Society,* 46 (12) 1075-1091, 1998.
15. Botteldooren, D. "Finite-Difference Time-Domain Simulations of Low-Frequency Room Acoustics Problems." *Journal of the Acoustical Society of America,* 8 (6). 3302-3309, 1995.
16. Lalor, N. "Statistical Energy Analysis and Its Use as an NVH Analysis Tool." *Sound & Vibration,*30 (1). 16-20, 1996.
17. CARA (Computer-Aided-Room-Acoustics) software, from www.cara.de, Version 2.2.
18. Bradley, J. S. "Sound Absorption of Gypsum Board Cavity Walls." *Journal of the Audio Engineering Society*, 45 (4), 253-259, 1997.
19. Cook, D., and Van Haverbeke, D. *"Trees and Shrubs for Noise Abatement."* Research Bulletin 246. The Forest Service, U.S. Department of Agriculture cooperating with the University of Nebraska College of Agriculture, 1971.
20. Probst, W., and Huber, B. "Calculation and Assessment of Traffic Noise Exposure." *Sound and Vibration,* 07/00 pp.16-20, 2000.
21. BASTIAN software, from DataKustic, Germany.
22. Kleiner, M., Dalenbäck, B.-I., and Svensson, P. "Auralization — An Overview." *Journal of the Audio Engineering Society,* 41 (11): 861-875, 1993.
23. Adapted from reference 1, Chapter 4, section 4, pp. 106, endnote [43].
24. CATT-Acoustics, from B.I. Dalenbäck
25. ODEON, from Ørsted DTU
26. EARS, from ADA (Acoustic Design Ahnert)
27. DIVA, from the DIVA Group
28. Wightman, F. L., and Kistler, D. J. "Headphone Simulation of Free-Field Listening. I: Stimulus Synthesis." *Journal of the Audio Engineering Society*, 85 (2). 858-867, 1989.

29. Wightman, F. L., and Kistler, D. J. “Headphone Simulation of Free-Field Listening. II: Psychological Validation.” *Journal of the Audio Engineering Society,* 85 (2). 868-878, 1989.
30. Camras, M. “Approach to Recreating a Sound Field.” *Journal of the Acoustical Society of America,* 43 (6). 1425-1431, 1968.
31. Ono, K., Komiyama, S., and Nakabayashi, K. “A Method of Reproducing Concert Hall Sounds by Loudspeaker Walls.” *Journal of the Acoustical Society of America,* 46 (11). 988-995, 1998.
32. Savioja, L. *Modeling Techniques for Virtual Acoustics.* Ph.D. Dissertation, 2000.
33. Dalenb**ä**ck Bengt-Inge and Str**ö**mberg, M. “Real time walkthrough auralization – The first year.” Proceedings of the Institute of Acoustics, 28 (2), 2007.
34. From Wikipedia reference on Ambisonics
35. Granier, E., Kleiner, M., Dalenbäck, B.-I., and Svensson, P. “Experimental Auralization of Car Audio Installations.” *Journal of the Audio Engineering Society,* 44 (10): 835-849, 1996.
36. Cadna A. “*Computer Program for the Calculation and Assessment of Outdoor Noise Immissions.”* User Document, DataKustic GmbH, 2000.

This page intentionally left blank

Part 2

Electronic Components

This page intentionally left blank

Chapter 10

Resistors, Capacitors, and Inductors

by Glen Ballou

10.1 Resistors 243
  10.1.1 Resistor Characteristics 243
  10.1.2 Combining Resistors 246
  10.1.3 Types of Resistors 247
10.2 Capacitors 251
  10.2.1 Time Constants 253
  10.2.2 Network Transfer Function 253
  10.2.3 Characteristics of Capacitors 253
  10.2.4 Types of Capacitors 255
    10.2.4.1 Film Capacitors 256
    10.2.4.2 Paper Foil-Filled Capacitors 257
    10.2.4.3 Mica Capacitors 257
    10.2.4.4 Ceramic Capacitors 257
    10.2.4.5 Electrolytic Capacitors 257
    10.2.4.6 Suppression Capacitors 261
    10.2.4.7 Supercapacitors 262
10.3 Inductors 263
  10.3.1 Types of Inductors 263
    10.3.1.1 Air Core Inductors 263
    10.3.1.2 Axial Inductor 263
    10.3.1.3 Bobbin Core Inductor 263
    10.3.1.4 Ceramic Core 263
    10.3.1.5 Epoxy-Coated Inductor 264
    10.3.1.6 Ferrite Core 264
    10.3.1.7 Laminated Cores 264
    10.3.1.8 Molded Inductor 264
    10.3.1.9 MPP Core 264
    10.3.1.10 Multilayer Inductor 264
    10.3.1.11 Phenolic Core 264
    10.3.1.12 Powdered Iron Core 264
    10.3.1.13 Radial Inductor 265
    10.3.1.14 Shielded Inductor 265
    10.3.1.15 Slug Core 265
    10.3.1.16 Tape Wound Core 265
    10.3.1.17 Toroidal Inductor 265
  10.3.2 Impedance Characteristics 265

10.3.3 Ferrite Beads 268
10.3.4 Skin Effect 268
10.3.5 Shielded Inductor 268
10.4 Impedance 268
10.5 Resonant Frequency 270
References 272

10.1 Resistors

Resistance is associated with the phenomenon of energy dissipation. In its simplest form, it is a measure of the opposition to the flow of current by a piece of electric material. Resistance dissipates energy in the form of heat; the best conductors have low resistance and produce little heat, whereas the poorest conductors have high resistance and produce the most heat. For example, if a current of 10 A flowed through a resistance of 1 Ω, the heat would be 100 W. If the same current flowed through 100 Ω, the heat would be 10,000 W, which is found with the equation

$$P = I^2R \tag{10-1}$$

where,
P is the power in watts,
I is the current in amperes,
R is the resistance in ohms.

In a pure resistance—i.e. one without inductance or capacitance—the voltage and current phase relationship remains the same. In this case the voltage drop across the resistor is

$$V = IR \tag{10-2}$$

where,
V is the voltage in volts,
I is the current in amperes,
R is the resistance in ohms.

All resistors have one by-product in common when put into a circuit, they produce heat because power is dissipated any time a voltage, V, is impressed across a resistance R. This power is calculated from Eq. 10-1 or

$$P = \frac{V^2}{R} \tag{10-3}$$

where,
P is the power in watts,
V is the voltage in volts,
R is the resistance in ohms.

Changing the voltage, while holding the resistance constant, changes the power by the square of the voltage. For instance, a voltage change from 10 V to 12 V increases the power 44%. Changing the voltage from 10 V to 20 V increases the power 400%.

Changing the current while holding the resistance constant has the same effect as a voltage change. Changing the current from 1 A to 1.2 A increases the power 44%, whereas changing from 1 A to 2 A increases the power 400%.

Changing the resistance while holding the voltage constant changes the power linearly. If the resistance is decreased from 1 kΩ to 800 Ω and the voltage remains the same, the power will increase 20%. If the resistance is increased from 500 Ω to 1 kΩ, the power will decrease 50%. Note that an increase in resistance causes a decrease in power.

Changing the resistance while holding the current constant is also a linear power change. In this example, increasing the resistance from 1 kΩ to 1.2 kΩ increases the power 20%, whereas increasing the resistance from 1 kΩ to 2 kΩ increases the power 100%.

It is important in sizing resistors to take into account changes in voltage or current. If the resistor remains constant and voltage is increased, current also increases linearly. This is determined by using Ohm's Law, Eq. 10-1 or 10-3.

Resistors can be fixed or variable, have tolerances from 0.5% to 20%, and power ranges from 0.1 W to hundreds of watts

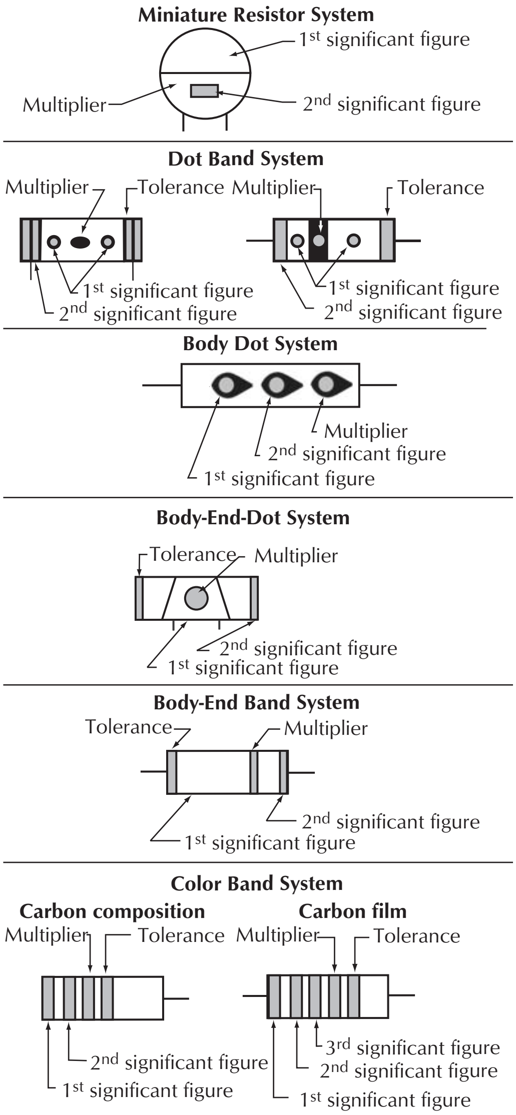
10.1.1 Resistor Characteristics

Resistors will change value as a result of applied voltage, power, ambient temperature, frequency change, mechanical shock, or humidity.

The values of the resistor are either printed on the resistor, as in power resistors, or are color coded on the resistor, Fig. 10-1. While many of the resistors in Fig. 10-1 are obsolete, they are still found in grandma's old radio you are asked to repair.

Voltage Coefficient. The *voltage coefficient* is the rate of change of resistance due to an applied voltage, given in percent parts per million per volt (%ppm/V). For most resistors the voltage coefficient is negative—that is, the resistance decreases as the voltage increases. However, some semiconductor devices increase in resistance with applied voltage. The voltage coefficient of very high valued carbon-film resistors is rather large and for wirewound resistors is usually negligible. *Varistors* are resistive devices designed to have a large voltage coefficient.

Temperature Coefficient of Resistance. The *temperature coefficient of resistance* (TCR) is the rate of change in resistance with ambient temperature, usually stated as parts per million per degree Celsius (ppm/°C). Many types of resistors increase in value as the temperature increases, while others, particularly hot-molded carbon types, have a maximum or minimum in their resistance


Miniature Resistor System
1st significant figure
Multiplier
2nd significant figure
Dot Band System
Multiplier
Tolerance
Multiplier
Tolerance
1st significant figure
2nd significant figure
1st significant figure
2nd significant figure
Body Dot System
Multiplier
2nd significant figure
1st significant figure
Body-End-Dot System
Tolerance
Multiplier
2nd significant figure
1st significant figure
Body-End Band System
Tolerance
Multiplier
2nd significant figure
1st significant figure
Color Band System
Carbon composition
Carbon film
Multiplier
Tolerance
Multiplier
Tolerance
2nd significant figure
1st significant figure
3rd significant figure
2nd significant figure
1st significant figure

Resistors with black body are composition, noninsulated.
Resistors with colored body are composition, insulated.
Wirewound resistors have the 1st color band double width.

Color	Digit	Multiplier	Tolerance	Failure Rate
Black	0	1	–	–
Brown	1	10	±1%	1.0
Red	2	100	±2%	0.1
Orange	3	1000	±3%	0.01
Yellow	4	10,000	±4%	0.001
Green	5	100,000	–	–
Blue	6	1,000,000	–	–
Violet	7	10,000,000	–	–
Gray	8	100,000,000	–	–
White	9	–	–	Solderable*
Gold		0.1	±5%	–
Silver		0.01	±10%	–
No Color			±20%	

Figure 10-1. Color codes for resistors.

curves that gives a zero temperature coefficient at some temperature. Metal film and wirewound types have temperature coefficient values of less than 100 ppm/°C. Thermistors are resistance devices designed to have a large temperature coefficient.

The percent temperature coefficient of resistance is

$$TCR = \frac{(R-r)100}{(T_R - T_T)R} \qquad (10\text{-}4)$$

where,
TCR is the temperature coefficient in percent per °C,
R is the resistance at reference temperature,
r is the resistance at test temperature,
T_R is the reference temperature in °C,
T_T is the test temperature in °C.

It is better to operate critical resistors with a limited temperature rise.

Noise. *Noise* is an unwanted voltage fluctuation generated within the resistor. The total noise of a resistor always includes Johnson noise, which depends only on resistance value and the temperature of the resistance element. Depending on the type of element and its construction, total noise may also include noise caused by current flow and by cracked bodies and loose end caps or leads. For adjustable resistors, noise is also caused by the jumping of the contact over turns of wire and by an imperfect electrical path between the contact and resistance element.

Hot-Spot Temperature. The *hot-spot temperature* is the maximum temperature measured on the resistor due to both internal heating and the ambient operating temperature. The maximum allowable hot-spot temperature is predicated on the thermal limits of the materials and the resistor design. The maximum hot-spot temperature may not be exceeded under normal operating conditions, so the wattage rating of the resistor must be lowered if it is operated at an ambient temperature higher than that at which the wattage rating was established. At zero dissipation, the maximum ambient temperature around the resistor may be its maximum hot-spot temperature. The ambient temperature for a resistor is affected by surrounding heat-producing devices. Resistors stacked together do not experience the actual ambient temperature surrounding the outside of the stack except under forced cooling conditions.

Carbon resistors should, at most, be warm to touch, 40°C (140°F), while wirewound or ceramic resistors are designed to operate at temperatures up to 140°C (284°F). Wherever power is dissipated, it is imperative that adequate ventilation is provided to eliminate thermal destruction of the resistor and surrounding components.

Power Coefficient. The *power coefficient* is the product of the temperature coefficient of resistance and the temperature rise per watt. It is given in percent per watt (%/W), and is the change in value resulting from applied power.

Ac Resistance. The *ac resistance* value changes with frequency because of the inherent inductance and capacitance of the resistor plus the skin effect, eddy current losses, and dielectric loss.

Ambient Temperature Effect. When operating a resistor in free air at high ambient temperature, the power capabilities must be derated, Fig. 10-2. Free air is operation of a resistor suspended by its terminals in free space and still air with a minimum clearance of one foot in all directions to the nearest object.

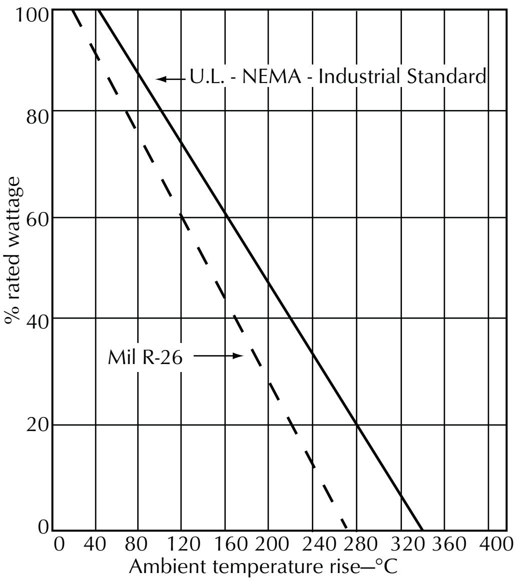


Figure 10-2. Resistor derating for elevated ambient temperature. Courtesy Ohmite Mfg. Co.

Grouping. Mounting a number of resistors in close proximity can cause excessive temperature rise requiring derating the power capabilities, Fig 10-3. The curves are for operation at maximum permissible hot spot temperature with spacing between the closest points of the resistors. Derating could be less if operated at less than permissible hot spot temperature.

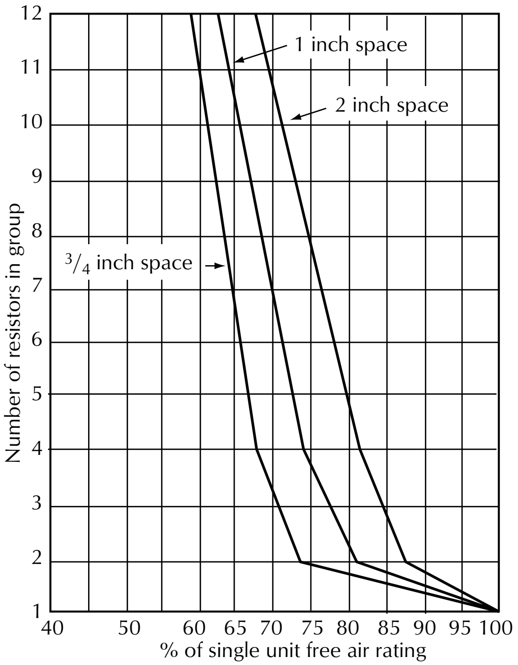
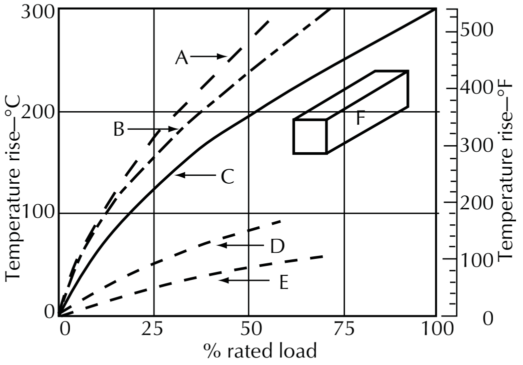


Figure 10-3. Power derating for grouping resistors. Courtesy Ohmite Mfg. Co.

Enclosure. Enclosures create a rise in temperature due to the surface area, size, shape, orientation, thickness, material and ventilation. Fig. 10-4 indicates the effects on a resistor enclosed in an unpainted steel sheet metal box, 0.32 inches thick without vents. Determining the derating is often by trial and error.



A. Resistor in $3^3/_8$ inch × $3^3/_8$ inch × 8 inch box.
B. Resistor in $5^{13}/_{16}$ inch × $5^{13}/_{16}$ inch × $12^3/_4$ inch box.
C. Resistor in free air.
D. Box temperature—small.
E. Box temperature—large.
F. Unpainted sheet metal box, 0.32 inch thick steel, no vents.

Figure 10-4. Effect of the size of an enclosure on a 500 W $^3/_4$ inch × $6^1/_2$ inch resistor. Courtesy of Ohmite Mfg. Co.

Forced Air Cooling. Resistors and components can be operated at higher than rated wattage with forced air cooling, Fig. 10-5. The volume of cooling air required to keep the resistor temperature within limits can be found with the equation

$$Volume\ of\ air = \frac{3170}{\Delta T} KW \qquad (10\text{-}5)$$

where,

Volume of air is in cubic feet per minute,

ΔT is the permissible temperature rise in degrees F,

KW is the power dissipated inside the enclosure in kilowatts.

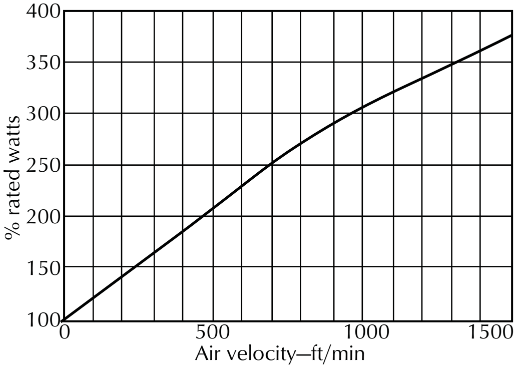


Figure 10-5. Percent of free air rating for a typical resistor cooled by forced air. Courtesy of Ohmite Mfg. Co.

Air density at high altitudes causes less heat to be dissipated by convection so more forced air would be required.

Pulse Operation. A resistor can usually be operated with a higher power in the pulse mode than in a continuous duty cycle. The actual increase allowed depends on the type of resistor. Fig. 10-6 is the percent of continuous duty rating for pulse operation for a wirewound resistor. Fig. 10-7 is the percent of continuous duty rating for pulse operation for typical NEMA duty cycles. Fig. 10-8 shows the percent of continuous duty rating for pulse operation of a 160 W vitreous enameled resistor.

10.1.2 Combining Resistors

Resistors can be combined is series or parallel or series/parallel.

Resistors in series. The total resistance of resistors connected in series is the summation of the resistors.

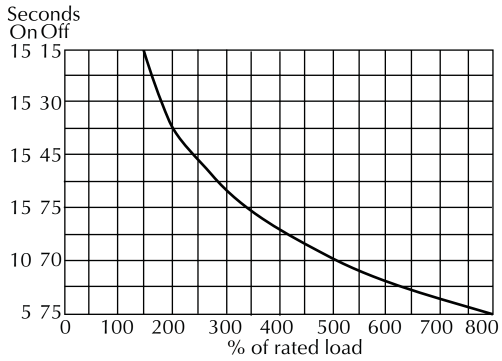


Figure 10-6. Effect of pulse operation on wirewound resistors. Courtesy of Ohmite Mfg. Co.

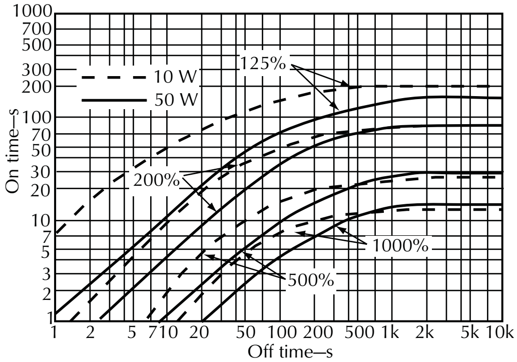


Figure 10-7. Percent of continuous duty rating for pulse operation of small and medium size vitreous enameled resistors. Courtesy of Ohmite Mfg. Co.

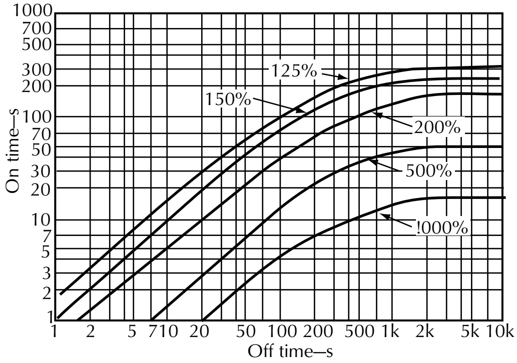


Figure 10-8. Percent of continuous duty rating for pulse operation of a 160 W vitreous enameled resistor. Courtesy of Ohmite Mfg. Co.

$$R_T = R_1 + R_2 + \ldots R_n \qquad (10\text{-}6)$$

The total resistance is always greater than the largest resistor.

Resistors in Parallel. The total resistance of resistors in parallel is

$$R_T = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \dots \frac{1}{R_n}} \qquad (10\text{-}7)$$

If two resistors are in parallel use:

$$R_T = \frac{R_1 \times R_2}{R_1 + R_2} \qquad (10\text{-}8)$$

When all of the resistors are equal, divide the value of one resistor by the number of resistors to determine the total resistance. The total resistance is always less than the smallest resistor.

To determine the value of one of the resistors when two are in parallel and the total resistance and one resistor in known, use

$$R_2 = \frac{R_T \times R_1}{R_1 - R_T} \qquad (10\text{-}9)$$

10.1.3 Types of Resistors

Every material that conducts electrical current has resistivity, which is defined as the resistance of a material to electric current. Resistivity is normally defined as the resistance, in ohms, of a 1 cm per side cube of the material measured from one surface of the cube to the opposite surface. The measurement is stated in ohms per centimeter cubed (Ω/cm^3). The inverse of resistivity is conductivity. Good conductors have low resistivity, and good insulators have high resistivity. Resistivity is important because it shows the difference between materials and their opposition to current, making it possible for resistor manufacturers to offer products with the same resistance but differing electrical, physical, mechanical, or thermal features.

Following is the resistivity of various materials:

Material	Resistivity
Aluminum	0.0000028
Copper	0.0000017
Nichrome	0.0001080
Carbon (varies)	0.0001850
Ceramic (typical)	100,000,000,000,000 or (10^{14})

Carbon-Composition Resistors. Carbon-composition resistors are the least expensive resistors and are widely used in circuits that are not critical to input noise and do not require tolerances better than ±5%.

The carbon-composition, hot-molded version is basically the same product it was more than 50 years ago. Both the hot- and cold-molded versions are made from a mixture of carbon and a clay binder. In some versions, the composition is applied to a ceramic core or armature, while in the inexpensive version, the composition is a monolithic rigid structure. Carbon-composition resistors may be from 1 Ω to many megohms and 0.1–4 W. The most common power rating is ¼ W and ½ W with resistance values from 2 Ω–22 MΩ.

Carbon-composition resistors can withstand higher surge currents than carbon-film resistors. Resistance values, however, are subject to change upon absorption of moisture and increase rapidly at temperatures much above 60°C (140°F). Noise also becomes a factor when carbon-composition resistors are used in audio and communication applications. A carbon-core resistor, for example, generates electrical noise that can reduce the readability of a signal or even mask it completely.

Carbon-Film Resistors. Carbon-film resistors are leaded ceramic cores with thin films of carbon applied. Carbon film resistors offer closer tolerances and better temperature coefficients than carbon composition resistors. Most characteristics are virtually identical for many general purpose, noncritical applications where high reliability, surge currents, or noise are not crucial factors.

Metal Film Resistors. Metal film resistors are discrete devices formed by depositing metal or metal oxide films on an insulated core. The metals are usually either nichrome sputtered on ceramic or tin oxide on ceramic or glass. Another method of production is to screen or paint powdered metal and powdered glass that is mixed in an ink or pastelike substance on a porous ceramic substrate. Firing or heating in an oven bonds the materials together. This type of resistor technology is called *cermet technology.*

Metal film resistors are most common in the 10 Ω to 1 MΩ range and ⅛ W to 1 W with tolerances of ±1%.

The TCR is in the ±100 ppm/°C range for all three technologies. Yet there are subtle differences:

- Cermet covers a wider resistance range and handles higher power than nichrome deposition.
- Nichrome is generally preferred over tin oxide in the upper and lower resistance ranges and can provide TCRs that are lower than 50 ppm/°C.
- Tin oxide is better able to stand higher power dissipation than nichrome.

Wirewound Resistors. *Wirewound resistors* have resistive wire wound on a central ceramic core. One of the oldest technologies, wirewounds provide the best known characteristics of high temperature stability and power handling ability. Nichrome, Manganin, and Evanohm are the three most widely used wires for wirewound resistors.

Wirewound resistors are usually in the 0.1 Ω–250 kΩ range. Tolerance is ±2% and TCR is ±10 ppm/°C.

Wirewound resistors are generally classed as power or instrument-grade products. Power wirewounds, capable of handling as much as 1500 W, are wound from uninsulated coarse wire to provide better heat dissipation. Common power ratings are 1.5 W, 3 W, 5 W, 8 W, 10 W, 20 W, 25 W, 50 W, 100 W, and 200 W.

Instrument-grade precision wirewound resistors are made from long lengths of finely insulated wire. After winding, they are usually coated with a ceramic material.

All wirewound resistors are classed as air-core inductors and the inductive reactance alters the high frequency resistive value. This problem is directly proportional with frequency. Special windings are useful to cancel reactance at audio frequencies. Because of the severity of the problem, these resistors cannot be used at high frequencies.

Noninductive Resistors. Non-inductive resistors are used for high frequency applications. This is accomplished by utilizing the Ayrton-Perry type of wiring, i.e. two windings connected in parallel and wound in opposite directions. This keeps the inductance and distributed capacitance at a minimum. Table 10-1 is a comparison of MEMCOR-TRUOHM type FR10, FR50, VL3 and VL5 resistors.

Resistor Networks. With the advent of printed circuit boards and integrated circuits, *resistor networks* became popular. The resistive network may be mounted in a single-in-line package (SIP) socket or a dual-in-line package (DIP) socket—the same as the ones used for integrated circuits. The most common resistor network has 14 or 16 pins and includes 7 or 8 individual resistors or 12 to 15 resistors with a common terminal. In most resistor networks the value of the resistors are the same. Networks may also have special value resistors and interconnections for a specific use, as shown in Fig. 10-9.

The individual resistors in a thick-film network can have a resistance value ranging from 10 Ω to 2.2 MΩ and are normally rated at 0.125 W per resistor. They have normal tolerances of ±2% or better and a temperature coefficient of resistance ±100 ppm/°C from −55°C to +125°C (–67°F to +257°F).

Table 10-1. Inductance Comparison of Standard and Non-Inductive Windings.

		Approximate Frequency Effect		
		Stock inductive winding	Non-inductive winding	
Type	Resistance (Ω)	L_S (μH)	L_S (μH)	C_P (μμF)
FR10 (10 W)	25	5.8	0.01	–
	100	11.0	0.16	–
	500	18.7	0.02	–
	1000	20.8	–	0.75
	5000	43.0	–	1.00
FR50 (50 W)	25	6.8	0.05	–
	100	>100.0	0.40	–
	500	>100.0	0.31	–
	1000	>100.0	–	1.10
	5000	>100.0	–	1.93
VL3 (3 W)	25	1.2	0.02	–
	100	1.6	0.07	–
	500	4.9	–	0.47
	1000	4.5	–	0.70
	5000	3.0	–	1.00
VL5 (5 W)	25	2.5	0.08	–
	100	5.6	0.14	–
	500	6.4	–	0.03
	1000	16.7	–	0.65
	5000	37.0	–	0.95

Courtesy Ohmite Mfg. Co.

Thin-film resistors are almost always specialized units and are packaged as DIPs or flatpacks. (Flatpacks are soldered into the circuit.) Thin-film networks use nickel chromium, tantalum nitride, and chromium cobalt vacuum depositions.

Variable Resistors. *Variable resistors* are ones whose value changes with light, temperature, or voltage or through mechanical means.

Photocells (Light-Sensitive Resistors). *Photocells* are used as off–on devices when a light beam is broken or as audio pickups for optical film tracks. In the latter, the sound track is either a variable density or variable area. Whichever, the film is between a focused light source and the photocell. As the light intensity on the photocell varies, the resistance varies.

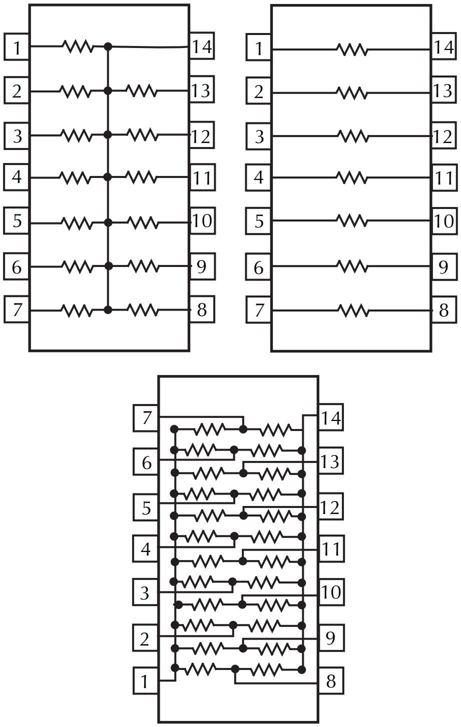


Figure 10-9. Various types of resistor networks.

Photocells are rated by specifying their resistance at low and high light levels. These typically vary from 600 Ω–110 kΩ (bright), and from 100 kΩ–200 MΩ (dark). Photocell power dissipation is between 0.005 W and 0.75 W.

Thermistors. *Thermistors*, thermal-sensitive resistors, may increase or decrease their resistance as temperature rises. If the coefficient of resistance is negative, the resistance decreases as the temperature increases; if positive, the resistance increases with an increase in temperature. Thermistors are specified by how their resistance changes for a 1°C change in temperature. They are also rated by their resistance at 25°C and by the ratio of resistance at 0°C and 50°C. Values vary from 2.5 Ω–1 MΩ at room temperature with power ratings from 0.1–1 W.

Thermistors are normally used as temperature-sensing devices or transducers. When used with a transistor, they can be used to control transistor current with a change in temperature. As the transistor heats up, the emitter-to-collector current increases. If the power supply voltage remains the same, the power dissipation in the transistor increases until it destroys itself through thermal runaway. The change in resistance due to temperature change of the thermistor placed in the base circuit of a transistor can be used to reduce base voltage and, therefore, reduce the transistor emitter to collector current. By properly matching the temperature coefficients of the two devices, the output current of the transistor can be held fairly constant with temperature change.

Varistors. *Varistors* (voltage-sensitive resistors) are voltage-dependent, nonlinear resistors which have symmetrical, sharp breakdown characteristics similar to back-to-back Zener diodes. They are designed for transient suppression in electrical circuits. The transients can result from the sudden release of previously stored energy—i.e., electromagnetic pulse (EMP)—or from extraneous sources beyond the control of the circuit designer, such as lightning surges. Certain semiconductors are most susceptible to transients. For example, LSI and VLSI circuits, which may have as many as 20,000 components in a 0.25 inch × 0.25 inch area, have damage thresholds below 100 μJ.

The *varistor* is mostly used to protect equipment from power-line surges by limiting the peak voltage across its terminals to a certain value. Above this voltage, the resistance drops, which in turn tends to reduce the terminal voltage. Voltage-variable resistors or varistors are specified by power dissipation (0.25 –1.5 W) and peak voltage (30–300 V).

Thermocouples. While not truly a resistor, thermocouples are used for temperature measurement. They operate via the Seebeck Effect which states that two dissimilar metals joined together at one end produce a voltage at the open ends that varies as the temperature at the junction varies. The voltage output increases as the temperature increases. Thermocouples are rugged, accurate, and have a wide temperature range. They don't require a exitation source and are highly responsive. Thermoouples are tip sensitive so they measure the temperature at a very small spot. Their output is very small (tens to hundreds of microvolts, and is nonlinear, requiring external linearization in the form of cold-junction compensation.

Never use copper wire to connect a thermocouple to the measureing device as that constitutes another thermocouple.

Resistance Temperature Detectors. RTDs are very accurate and stable. Most are made of platinum wire wound around a small ceramic tube. They can be thermally shocked by going from 100°C to –195°C 50 times with a resulting error less than 0.02°C.

RTDs feature a low resistance-value change to temperature (0.1 Ω/1°C. RTDs can self heat, causing inaccurate readings, therefore the current through the unit should be kept to 1 mA or less. Self heating can also be controlled by using a 10% duty cycle rather than constant bias or by using an extremely low bias which can reduce the SNR. The connection leads may cause errors if they are long due to the wire resistance.

Potentiometers and Rheostats. The resistance of *potentiometers* (pots), and *rheostats* is varied by mechanically varying the size of the resistor. They are normally three terminal devices, two ends and one wiper, Fig. 10-10. By varying the position of the wiper, the resistance between either end and the wiper changes. Potentiometers may be wirewound or nonwirewound. The nonwirewound resistors usually have either a carbon or a conductive plastic coating. Potentiometers or pots may be 300° single turn or multiple turn, the most common being 1080° three turn and 3600° ten turn.

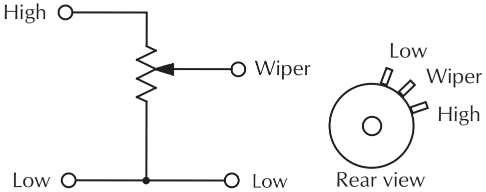


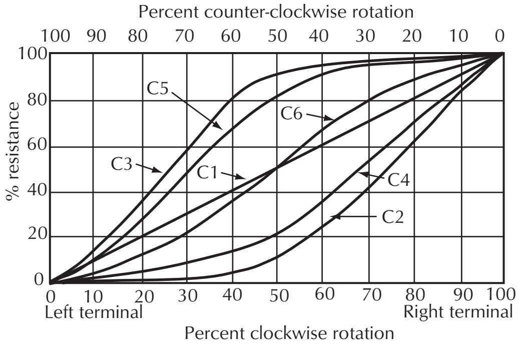
Figure 10-10. Three terminal potentiometer.

Wirewound pots offer TCRs of ±50 ppm/°C and tolerances of ±5%. Resistive values are typically 10 Ω–100 kΩ, with power ratings from 1 W to 200 W.

Carbon pots have TCRs of ±400 ppm/°C to ±800 ppm/°C and tolerances of ±20%. The resistive range spans 50 Ω–2 MΩ, and power ratings are generally less than 0.5 W.

Potentiometers may be either linear or nonlinear, as shown in Fig. 10-11. The most common nonlinear pots are counterclockwise semilog and clockwise semilog. The counterclockwise semilog pot is also called an audio taper pot because when used as a volume control, it follows the human hearing equal loudness curve. If a linear pot is used as a simple volume control, only about the first 20% of the pot rotation would control the usable volume of the sound system. By using an audio taper pot as in Fig. 10-11 curve C2, the entire pot is used. Note there is only a 10%–20% change in resistance value between the common and wiper when the pot is 50% rotated.

Potentiometers are also produced with various taps that are often used in conjunction with *loudness* controls.



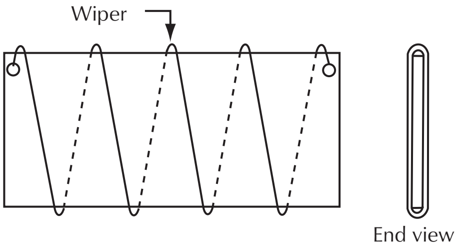
C1. Linear taper, general-purpose control for television picture adjustments. Resistance proportional to shaft rotation.
C2. Left-hand semilog taper for volume and tone controls. 10% of resistance at 50% rotation.
C3. Right-hand semilog taper, reverse of C2. 90% of resistance at 50% of rotation.
C4. Modified left-hand semilog taper for volume and tone controls. 20% of resistance at 50% of rotation.
C5. Modified right-hand semilog taper, reverse of C4. 80% of resistance at 50% of rotation.
C6. Symmetrical straight-line taper with slow resistance change at either end. Used principally as tone control or balance control.

Figure 10-11. Tapers for six standard potentiometers in resistivity versus rotation.

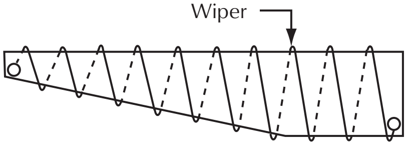
Potentiometers also come in combinations of two or more units controlled by a single control shaft or controlled individually by concentric shafts. Switches with various contact configurations can also be assembled to single or ganged potentiometers and arranged for actuation during the first few degrees of shaft rotation.

A *wirewound potentiometer* is made by winding resistance wire around a thin insulated card, Fig. 10-12A. After winding, the card is formed into a circle and fitted around a form. The card may be tapered, Fig. 10-12B, to permit various rates of change of resistance as shown in Fig 10-11. The wiper presses along the wire on the edge of the card.

Contact Resistance. Noisy potentiometers have been a problem that has plagued audio circuits for years. Although pots have become better in tolerance and construction, noise is still the culprit that forces pots to be replaced. Noise is usually caused by dirt or, in the case of wirewound potentiometers, oxidation. Many circuits have gone up in smoke because bias-adjusting resistors, which are wirewound for good TCR, oxidize and the contact resistance increases to a point where it is more than the value of the pot. This problem is most



A. Fixed card wirewound resistor.



B. Tapered card wirewound resistor.

Figure 10-12. Construction of a wirewound resistor.

noticeable when trying to adjust a bias voltage with an old oxidized pot.

Sometimes the pot can be cleaned by spraying it with a contact cleaner or silicone and then vigorously rotating it. Usually, however, it is best to replace it because anything else is only temporary.

Any dc voltage present on the pot is also a source of noise. Such voltage is often produced by leaky coupling capacitors at the input connector or output circuit of the wiper, allowing dc voltage to appear at the wiper contact. If there is a resistance between the resistor and the wiper, the dc current flowing through the wiper contact to the output stage will create a voltage drop. Because the wiper is moving, the contact resistance constantly changes creating what looks like a varying ac voltage. Using Fig. 10-13, the value at V_{Load}, whether ac or dc, can be calculated with Eqs. 10-10 and 10-11. If the wiper resistance is 0—i.e., a perfect pot—the output voltage V_{Load} is

$$V_{Load} = V_1\left(\frac{R_y}{R_1 + R_y}\right) \tag{10-10}$$

where,

$$R_y = \frac{R_2 R_{Load}}{R_2 + R_{Load}}$$

If a pot wiper has a high resistance, R_w, the output voltage V_{Load} is

$$V_{Load} = V_w\left(\frac{R_{Load}}{R_w + R_{Load}}\right) \tag{10-11}$$

where,

$$V_w = V_1\left(\frac{R_2(R_w + R_{Load})}{R_2 + R_w + R_{Load}}\right).$$

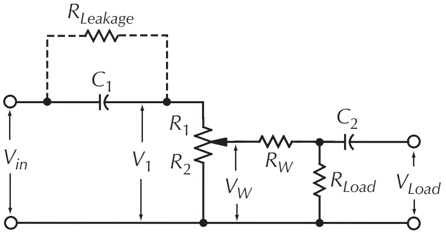


Figure 10-13. Effects of wiper noise on potentiometer output.

10.2 Capacitors

Capacitors are used for both dc and ac applications. In dc circuits they are used to store and release energy such as filtering power supplies and for providing on demand, a single high voltage pulse of current.

In ac circuits capacitors are used to block dc, allowing only ac to pass, bypassing ac frequencies, or discriminating between higher and lower ac frequencies. In a circuit with a pure capacitor, the current will lead the voltage by 90°.

The value of a capacitor is normally written on the capacitor and the sound engineer is only required to determine their effect in the circuit.

Where capacitors are connected in series with each other, the total capacitance is

$$C_T = \frac{1}{\frac{1}{C_1} + \frac{1}{C_2} \cdots + \frac{1}{C_n}} \tag{10-12}$$

and is always less than the value of the smallest capacitor.

When connected in parallel, the total capacitance is

$$C_T = C_1 + C_2 \ldots + C_n \tag{10-13}$$

and is always larger than the largest capacitor.

When a dc voltage is applied across a group of capacitors connected in series, the voltage drop across the combination is equal to the applied voltage. The drop across each individual capacitor is inversely proportional to its capacitance, and assuming each capacitor has an infinitely large effective shunt resistance, can be calculated by the equation

$$V_C = V_A\left(\frac{C_X}{C_T}\right) \tag{10-14}$$

where,

V_C is the voltage across the individual capacitor in the series (C_1, C_2,$\cdots$ C_n) in volts,

V_A is the applied voltage in volts,
C_X is the capacitance of the individual capacitor under consideration in farads,
C_T is the sum of all of the capacitors in series.

When used in an ac circuit, the capacitive reactance, or the impedance the capacitor injects into the circuit, is important to know and is found with the equation:

$$X_C = \frac{1}{2\pi f C} \quad (10\text{-}15)$$

where,
X_C is the capacitive reactance in ohms,
f is the frequency in hertz,
C is the capacitance in farads.

To determine the impedance of circuits with resistance, capacitance, and inductance, see Section 10.4.

Capacitance is the concept of energy storage in an electric field. If a potential difference is found between two points, an electric field exists. The electric field is the result of the separation of unlike charges, therefore, the strength of the field will depend on the amounts of the charges and their separator. The amount of work necessary to move an additional charge from one point to the other will depend on the force required and therefore upon the amount of charge previously moved. In a capacitor, the charge is restricted to the area, shape, and spacing of the capacitor electrodes, sometimes known as *plates*, as well as the property of the material separating the plates.

When electrical current flows into a capacitor, a force is established between two parallel plates separated by a dielectric. This energy is stored and remains even after the input current flow ceases. Connecting a conductor across the capacitor provides a plate-to-plate path by which the charged capacitor can regain electron balance, that is, discharge its stored energy. This conductor can be a resistor, hard wire, or even air. The value of a parallel plate capacitor can be found with the equation

$$C = \frac{x\varepsilon[(N-1)A]}{d} \times 10^{-13} \quad (10\text{-}16)$$

where,
C is the capacitance in farads,
x is 0.0885 when A and d are in cm, and 0.225 when A and d are in inches,
ε is the dielectric constant of the insulation,
N is the number of plates,
A is the area of the plates,
d is the spacing between the plates.

The work necessary to transport a unit charge from one plate to the other is

$$e = kg \quad (10\text{-}17)$$

where,
e is the volts expressing energy per unit charge,
k is the proportionality factor between the work necessary to carry a unit charge between the two plates and the charge already transported and is equal to $1/C$ where C is the capacitance in farads,
g is the coulombs of charge already transported.

The value of a capacitor can now be calculated from the equation

$$C = \frac{q}{e} \quad (10\text{-}18)$$

where,
q is the charge in coulombs,
e is found with Eq. 10-17.

The energy stored in a capacitor is found with the equation

$$W = \frac{CV^2}{2} \quad (10\text{-}19)$$

where,
W is the energy in joules,
C is the capacitance in farads,
V is the applied voltage in volts.

Dielectric Constant (K). The dielectric constant is the property of a given material that determines the amount of electrostatic energy that may be stored in that material per unit volume for a given voltage. The value of K expresses the ratio of a capacitor in a vacuum to one using a given dielectric. The K of air is 1 and is the reference unit employed for expressing K of other materials. If K of the capacitor is increased or decreased, the capacitance will increase or decrease respectively if other quantities and physical dimensions are kept constant. Table 10-2 is a listing of K for various materials.

Table 10-2. Comparison of Capacitor Dielectric Constants

Dielectric	K (Dielectric Constant)
Air or vacuum	1.0
Paper	2.0–6.0
Plastic	2.1–6.0
Mineral oil	2.2–2.3
Silicone oil	2.7–2.8

Table 10-2. Comparison of Capacitor Dielectric Constants (Continued)

Dielectric	K (Dielectric Constant)
Quartz	3.8–4.4
Glass	4.8–8.0
Porcelain	5.1–5.9
Mica	5.4–8.7
Aluminum oxide	8.4
Tantalum pentoxide	26.0
Ceramic	12.0–400,000

The dielectric constant of materials is generally affected by both temperature and frequency, except for quartz, Styrofoam, and Teflon, whose dielectric constants remain essentially constant. Small differences in the composition of a given material will also affect the dielectric constant.

Force. The equation for calculating the force of attraction between the two plates is

$$F = \frac{AV^2}{K(1504S)^2} \qquad (10\text{-}20)$$

where,

F is the attractive force in dynes,
A is the area of one plate in square centimeters,
V is the potential energy difference in volts,
K is the dielectric constant,
S is the separation between the plates in centimeters.

10.2.1 Time Constants

When a dc voltage is impressed across a capacitor, a time (t) is required to charge the capacitor to a voltage. This is determined with the equation:

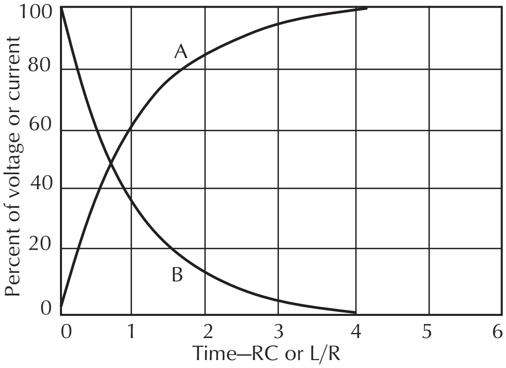
$$t = RC \qquad (10\text{-}21)$$

where,

t is the time in seconds,
R is the resistance in ohms,
C is the capacitance in farads.

In a circuit consisting of only resistance and capacitance, the time constant t is defined as the time it takes to charge the capacitor to 63.2% of the maximum voltage. During the next time constant, the capacitor is charged or the current builds up to 63.2% of the remaining difference of full value, or to 86.5% of the full value. Theoretically, the charge on a capacitor or the current through a coil can never actually reach 100% but is considered to be 100% after five time constants have passed. When the voltage is removed, the capacitor discharges and the current decays 63.2% per time constant to zero.

These two factors are shown graphically in Fig. 10-14. Curve A shows the voltage across a capacitor when charging. Curve B shows the capacitor voltage when discharging. It is also the voltage across the resistor on charge or discharge.



A. Voltage across *C* when charging.
B. Voltage across *C* when discharging.

Figure 10-14. Universal time graph.

10.2.2 Network Transfer Function

Network transfer functions are the ratio of the output to input voltage (generally a complex number) for a given type of network containing resistive and reactive elements. The transfer functions for networks consisting of resistance and capacitance are given in Fig. 10-15. The expressions for the transfer functions of the networks are:

A is $j\omega$ or $j2\pi f$,
B is RC,
C is R_1C_1,
D is R_2C_2,
n is a positive multiplier,
f is the frequency in hertz,
C is in farads,
R is in ohms.

10.2.3 Characteristics of Capacitors

The operating characteristics of a capacitor determine what it was designed for and therefore where it is best used.

Capacitance (*C*). The *capacitance* of a capacitor is normally expressed in microfarads (µF or 10^{-6} farads)

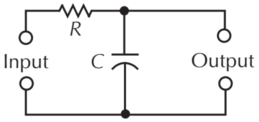
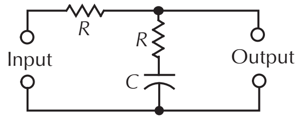
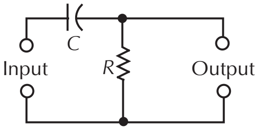
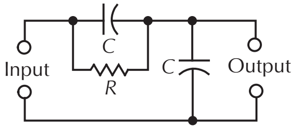
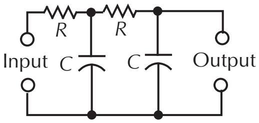
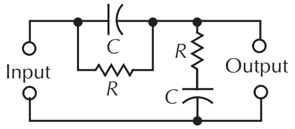
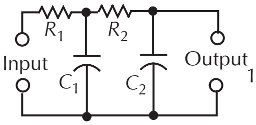
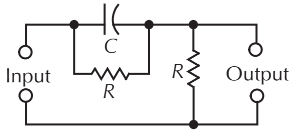
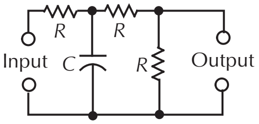
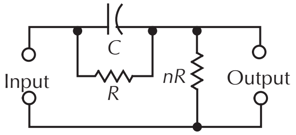
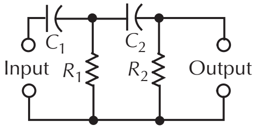
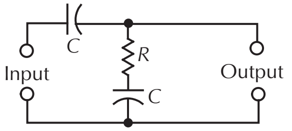
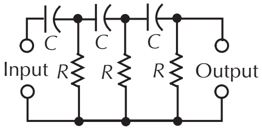
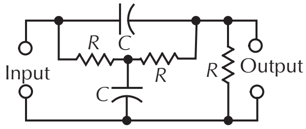
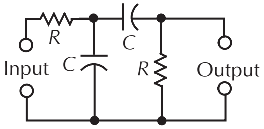
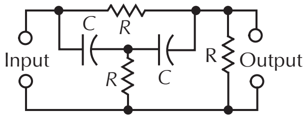
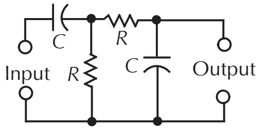
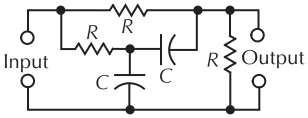
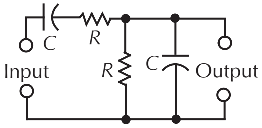
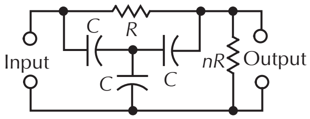
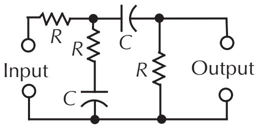
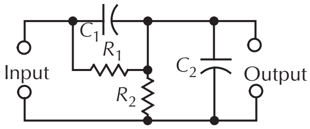
Network	Transfer function	Network	Transfer function
	$\frac{1}{1 + AB}$		$\frac{1 + AB}{1 + 2AB}$
	$\frac{AB}{1 + AB}$		$\frac{1 + AB}{1 + 2AB}$
	$\frac{1}{1 + 3AB + A^2B^2}$		$\frac{(1 + AB)^2}{1 + 3AB + A^2B^2}$
	$\frac{1}{1 + (C - D + R_1C_2)A + CDA^2}$		$\frac{1 + AB}{2 + AB}$
	$\frac{1}{3 + 2AB}$		$\frac{n(1 + AB)}{(1 = n) + nAB}$
	$\frac{CDA^2}{1 + (C - D + R_1C_2)A + CDA^2}$		$\frac{1 + AB}{2 + AB}$
	$\frac{A^3B^3}{1 + 5AB + 6A^2B^2 + A^3B^3}$		$\frac{1 + AB}{3 + AB}$
	$\frac{AB}{1 + 3AB + A^2B^2}$		$\frac{(1 + AB)^2}{2 + 5AB + A^2B^2}$
	$\frac{AB}{1 + 3AB + A^2B^2}$		$\frac{1 + 3AB}{2 + 5AB + A^2B^2}$
	$\frac{AB}{1 + 3AB + A^2B^2}$		$\frac{n(3 + AB)}{3(1 + n) + 2nAB}$
	$\frac{AB}{1 + 3AB}$		$\frac{R_2(1 + AC)}{(R_1 + R_2) + (R_1D + R_2C)A}$

Figure 10-15. Resistance-capacitance network transfer functions.

or picofarads (pF or 10^{-12} farads) with a stated accuracy or tolerance. Tolerance is expressed as plus or minus a certain percentage of the nominal or nameplate value. Another tolerance rating is GMV (guaranteed minimum value), sometimes referred to as MRV (minimum rated value). The capacitance will never be less than the marked value when used under specified operating conditions but the capacitance could be more than the named value.

Equivalent Series Resistance (ESR). All capacitors have an *equivalent series resistance* expressed in ohms or milliohms. This loss comes from lead resistance, termination losses, and dissipation in the dielectric material.

Equivalent Series Inductance (ESL). The *equivalent series inductance* can be useful or detrimental. It does reduce the high-frequency performance of the capacitor. However, it can be used in conjunction with the capacitors capacitance to form a resonant circuit.

Dielectric Absorption (DA). *Dielectric absorption* is a reluctance on the part of the dielectric to give up stored electrons when the capacitor is discharged. If a capacitor is discharged through a resistance, and the resistance is removed, the electrons that remained in the dielectric will reconvene on the electrode, causing a voltage to appear across the capacitor. This is also called *memory*.

When an ac signal, such as sound, with its high rate of attack is impressed across the capacitor, time is required for the capacitor to follow the signal because the free electrons in the dielectric move slowly. The result is compressed signal. The procedure for testing DA calls for a 5 min capacitor charging time, a 5 s discharge, then a 1 min open circuit, after which the recovery voltage is read. The percentage of DA is defined as the ratio of recovery to charging voltage times 100.

Insulation Resistance. *Insulation resistance* is basically the resistance of the dielectric material, and determines the period of time a capacitor, once charged with a dc voltage, will hold its charge by a specified percentage. The insulation resistance is generally very high. In electrolytic capacitors, the leakage current should not exceed

$$I_L = 0.04C + 0.30 \tag{10-22}$$

where,
I_L is the leakage current in microamperes,
C is the capacitance in microfarads.

Maximum Working Voltage. All capacitors have a *maximum working voltage* that should not be exceeded. The capacitors working voltage is a combination of the dc value plus the peak ac value that may be applied during operation. For instance, if a capacitor has 10 V_{dc} applied to it, and an ac voltage of 10 V_{rms} or 17 V_{peak} is applied, the capacitor will have to be capable of withstanding 27 V.

Quality Factor (*Q*). The *quality factor* of a capacitor is the ratio of the capacitors reactance to its resistance at a specified frequency. *Q* is found by the equation

$$Q = \frac{1}{2\pi fCR} \tag{10-23}$$

where,
f is the frequency in hertz,
C is the value of capacitance in farads,
R is the internal resistance in ohms.

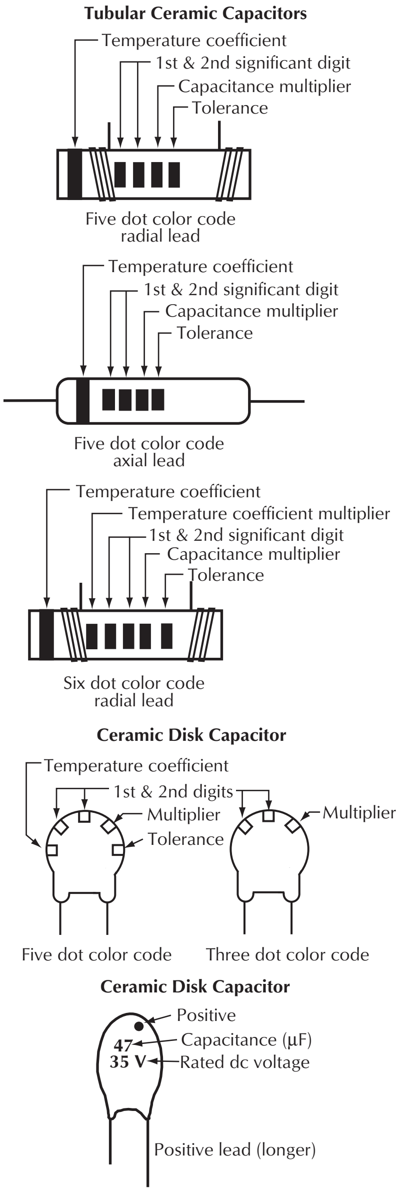
Dissipation Factor (DF). The *dissipation factor* is the ratio of the effective series resistance of a capacitor to its reactance at a specified frequency and is given in percent. It is also the reciprocal of *Q*. It is, therefore, a similar indication of power loss within the capacitor and, in general, should be as low as possible.

Power Factor (PF). The *power factor* represents the fraction of input volt-amperes or power dissipated in the capacitor dielectric and is virtually independent of the capacitance, applied voltage, and frequency. *PF* is the preferred measurement in describing capacitive losses in ac circuits.

10.2.4 Types of Capacitors

The uses made of capacitors become more varied and more specialized each year. They are used to filter, tune, couple, block dc, pass ac, shift phase, bypass, feed through, compensate, store energy, isolate, suppress noise, and start motors, among other things. While doing this, they frequently have to withstand adverse conditions such as shock, vibration, salt spray, extreme temperatures, high altitude, high humidity, and radiation. They must also be small, lightweight, and reliable.

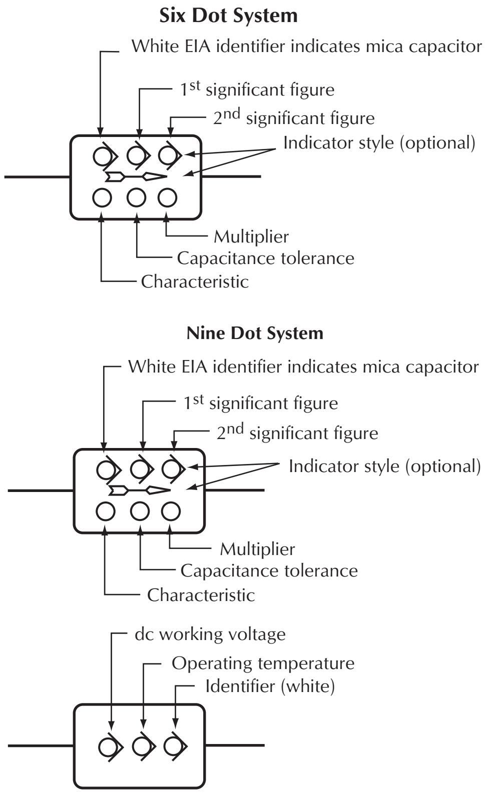
Capacitors are grouped according to their dielectric material and mechanical configuration. Because they may be hardwired or mounted on circuit boards, capacitors come with leads on one end, two ends, or they may be mounted in a dual-in-line (DIP) or single in-line (SIP) package. Figs. 10-16 and 10-17 show the various types of capacitors, their characteristics, and their color codes.


Tubular Ceramic Capacitors
Temperature coefficient
1st & 2nd significant digit
Capacitance multiplier
Tolerance
Five dot color code radial lead
Temperature coefficient
1st & 2nd significant digit
Capacitance multiplier
Tolerance
Five dot color code axial lead
Temperature coefficient
Temperature coefficient multiplier
1st & 2nd significant digit
Capacitance multiplier
Tolerance
Six dot color code radial lead
Ceramic Disk Capacitor
Temperature coefficient
1st & 2nd digits
Multiplier
Tolerance
Multiplier
Five dot color code
Three dot color code
Ceramic Disk Capacitor
Positive
47
35 V
Capacitance (μF)
Rated dc voltage
Positive lead (longer)

Capacitance in picofarads / Temperature coefficient

Color	Digits 1st	Digits 2nd	Multiplier	Tolerance 10 pF	Tolerance >10 pF	5 dot pp°C	6 dot Sig. Fig.	6 dot Multiplier
Black	0	0	1	±2.0 pF	±20%	0	0.0	−1
Brown	1	1	10	±0.1 pF	±1%	−33		−10
Red	2	2	100		±2%	−75	1.0	−100
Orange	3	3	1000		±3%	−150	1.5	−1000
Yellow	4	4				−230	2.0	−10,000
Green	5	5		±0.5 pF	±5%	−300	3.3	+1
Blue	6	6				−470	4.7	+10
Violet	7	7				−750	7.5	+100
Gray	8	8	0.01	±0.25 pF		−150 to −1500		+1000
White	9	9	0.1	±1.0 pF	10%	+100 to −75		+10,000
Gold	—	—						
Silver	—	—						

Figure 10-16. Color codes for tubular and disk ceramic capacitors.


Six Dot System
White EIA identifier indicates mica capacitor
1st significant figure
2nd significant figure
Indicator style (optional)
Multiplier
Capacitance tolerance
Characteristic
Nine Dot System
White EIA identifier indicates mica capacitor
1st significant figure
2nd significant figure
Indicator style (optional)
Multiplier
Capacitance tolerance
Characteristic
dc working voltage
Operating temperature
Identifier (white)

Mica Capacitor Characteristics

Characteristic	Temperature coefficient of capacitance (ppm/°C)	Maximum capacitance drift
B	Not specified	Not specified
C	±200	±0.5% +0.5pF
D	±100	±0.3% +0.1pF
E	−20–(+100)	±0.3% +0.1pF
F	0 to +70	±0.5% +0.1pF

Color	Char.	Digits 1st	Digits 2nd	Multiplier	Tolerance	dc working voltage	Operating temperature range
Black		0	0	1	±20%	100	
Brown	B	1	1	10	±1%		−55°C to +85°C
Red	C	2	2	100	±2%	300	
Orange	D	3	3	1000			−55°C to +125°C
Yellow	E	4	4	10,000		500	
Green	F	5	5		±5%		
Blue		6	6				
Violet		7	7				
Gray		8	8				
White		9	9				
Gold		–	–	0.1	±0.5%	1000	
Silver		–	–	0.01	±10%		

Figure 10-17. Color codes for mica capacitors.

10.2.4.1 Film Capacitors

Film capacitors consist of alternate layers of metal foil, and one or more layers of a flexible plastic insulating material (dielectric) in ribbon form rolled and encapsulated.

10.2.4.2 Paper Foil-Filled Capacitors

Paper foil-filled capacitors consist of alternate layers of aluminum foil and paper rolled together. The paper may be saturated with oil and the assembly mounted in an oil-filled, hermetically sealed metal case. These capacitors are often used as motor capacitors and are rated at 60 Hz.

10.2.4.3 Mica Capacitors

Two types of *mica capacitors* are in use. In one type, alternate layers of metal foil and mica insulation, are stacked together and encapsulated. In the silvered-mica type, a silver electrode is screened on the mica insulators that are then assembled and encapsulated. Mica capacitors have small capacitance values and are usually used in high frequency circuits.

10.2.4.4 Ceramic Capacitors

Ceramic capacitors are the most popular capacitors for bypass and coupling applications because of their variety of sizes, shapes, and ratings.

Ceramic capacitors also come with a variety of K values or dielectric constant. The higher the K value, the smaller the size of the capacitor. However, high K-value capacitors are less stable. High-K capacitors have a dielectric constant over 3000, are very small, and have values between 0.001 µF to several microfarads.

When temperature stability is important, capacitors with a K in the 10–200 region are required. If a high Q capacitor is also required, the capacitor will be physically larger. Ceramic capacitors can be made with a zero capacitance/temperature change. These are called *negative-positive-zero* (NPO). They come in a capacitance range of 1.0 pF–0.033 µF.

A temperature-compensated capacitor with a designation of N750 is used when temperature compensation is required. The 750 indicates that the capacitance will decrease at a rate of 750 ppm/°C with a temperature rise or the capacitance value will decrease 1.5% for a 20°C (68°F) temperature increase. N750 capacitors come in values between 4.0 pF and 680 pF.

10.2.4.5 Electrolytic Capacitors

The first electrolytic capacitor was made in Germany in about 1895 although its principle was discovered some 25 years earlier. It was not until the late 1920s when power supplies replaced batteries in radio receivers, that aluminum electrolytics were used in any quantities. The first electrolytics contained liquid electrolytes. These *wet* units disappeared during the late 1930s when the *dry gel* types took over.

Electrolytic capacitors are still not perfect. Low temperatures reduce performance and can even freeze electrolytes, while high temperatures can dry them out and the electrolytes themselves can leak and corrode the equipment. Also, repeated surges over the rated working voltage, excessive ripple currents, and high operating temperature reduce performance and shorten capacitor life. Even with their faults, electrolytic capacitors account for one-third of the total dollars spent on capacitors, probably because they provide high capacitance in small volume at a relatively low cost per microfarad-volt.

During the past few years, many new and important developments have occurred. Process controls have improved performance. Better seals have assured longer life, improved etching has given a tenfold increase in volume efficiencies, and leakage characteristics have improved one hundredfold.

Basic to the construction of electrolytic capacitors is the electrochemical formation of an oxide film on a metal surface. Intimate contact is made with this oxide film by means of another electrically conductive material. The metal on which the oxide film is formed serves as the anode or positive terminal of the capacitor; the oxide film is the dielectric, and the cathode or negative terminal is either a conducting liquid or a gel. The most commonly used basic materials are aluminum and tantalum.

Aluminum Electrolytic Capacitors. *Aluminum electrolytic capacitors* use aluminum as the base material. The surface is often etched to increase the surface area as much as 100 times that of unetched foil, resulting in higher capacitance in the same volume.

The type of etch pattern and the degree to which the surface area is increased involve many carefully controlled variables. If a fine etch pattern is desired to achieve high capacitance per unit area of foil for low voltage devices, the level of current density and time the foil is exposed to the etching solution will be far different from that required for a coarse etch pattern. The foil is then electrochemically treated to form a layer of aluminum oxide on its surface. Time and current density determine the amount of power consumed in the process. The oxide film dielectric is thin, usually about 15 Å/V. When formed on a high purity aluminum foil, it has a dielectric constant between 7 and 10 and an equivalent dielectric strength of 25 million volts per inch (25×10^6 V/inch).

The thickness of the oxide coating dielectric is determined by the voltage used to form it. The working voltage of the capacitor is somewhat less than this formation voltage. Thin films result in low voltage, high capacitance units; thicker films produce higher voltage, lower capacitance units for a given case size.

As a capacitor section is wound, a system of paper spacers is put in place to separate the foils. This prevents the possibility of direct shorts between anode and cathode foils that might result because of rough surfaces or jagged edges on either foil. The spacer material also absorbs the electrolyte with which the capacitor is impregnated, and thus assures uniform and intimate contact with all of the surface eccentricities of the etched anode foil throughout the life of the capacitor. The cathode foil serves only as an electrical connection to the electrolyte which is in fact the true cathode of the electrolytic capacitor.

The electrolyte commonly used in aluminum electrolytic capacitors is an ionogen that is dissolved in and reacts with glycol to form a pastelike mass of medium resistivity. This is normally supported in a carrier of high purity craft or hemp paper. In addition to the glycol electrolyte, low resistivity nonaqueous electrolytes are used to obtain a lower ESR and wider operating temperatures.

The foil-spacer-foil combination is wound into a cylinder, inserted into a suitable container, impregnated, and sealed.

- **Electrical Characteristics.** The equivalent circuit of an electrolytic capacitor is shown in Fig. 10-18. A and B are the capacitor terminals. The shunt resistance, R_s, in parallel with the effective capacitance, C, accounts for the dc leakage current through the capacitor. Heat is generated in the ESR if there is ripple current and heat is generated in the shunt resistance by the voltage. In an aluminum electrolytic capacitor, the ESR is due mainly to the spacer-electrolyte-oxide system. Generally it varies only slightly except at low temperatures where it increases greatly. L is the self-inductance of the capacitor caused by terminals, electrodes, and geometry.
- **Impedance.** The *impedance* of a capacitor is frequency dependent, as shown in Fig. 10-19. Here, ESR is the equivalent series resistance, X_C is the capacitive reactance, X_L is the inductive reactance, and Z is the impedance. The initial downward slope is a result of the capacitive reactance. The trough (lowest impedance) portion of the curve is almost totally resistive, and the rising upper or higher frequency portion of the curve is due to the capacitor's self-inductance. If the ESR were plotted separately, it would show a small ESR decrease with frequency to about 5–10 kHz, and then remain relatively constant throughout the remainder of the frequency range.

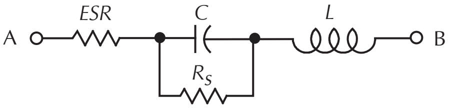


Figure 10-18. Simplified equivalent circuit of an electrolytic capacitor.

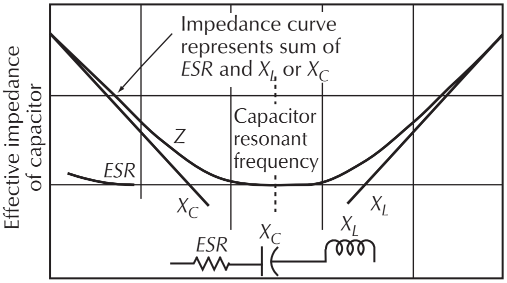


Figure 10-19. Impedance characteristics of a capacitor.

- **Leakage Current.** *Leakage current* in an electrolytic capacitor is the direct current that passes through a capacitor when a correctly polarized dc voltage is applied to its terminals. This current is proportional to temperature and becomes increasingly important when capacitors are used at elevated ambient temperatures. Imperfections in the oxide dielectric film cause high leakage currents. Leakage current decreases slowly after a voltage is applied and usually reaches *steady-state* conditions after 10 minutes.

  If a capacitor is connected with its polarity backward, the oxide film is forward biased and offers very little resistance to current flow, resulting in high current, which, if left unchecked, will cause overheating and self destruction of the capacitor.

  The total heat generated within a capacitor is the sum of the heat created by the I^2R losses in the ESR and that created by the $I_{Leakage} \times V_{applied}$.
- **Ac Ripple Current.** The *ac ripple current* rating is one of the most important factors in filter applications, because excessive current produces a greater than permissible temperature rise, shortening capacitor life. The maximum permissible rms ripple

current for any capacitor is limited by the temperature within the capacitor and the rate of heat dissipation from the capacitor. Lower ESR and longer cans or enclosures increase the ripple current rating.

- **Reverse Voltage.** Aluminum electrolytic capacitors can withstand a *reverse voltage* of up to 1.5 V without noticeable effect from its operating characteristics. Higher reverse voltages, when applied over extended periods, will lead to some loss of capacitance. Excess reverse voltages applied for short periods will cause some change in capacitance but may not lead to capacitor failure during the reverse voltage application or during subsequent operation in the normal polarity direction.

A major use of large value capacitors is for filtering in dc power supplies. After a capacitor is fully charged, when the rectifier conduction decreases, the capacitor discharges into the load until the next half cycle, Fig. 10-20. Then on the next cycle the capacitor recharges again to the peak voltage. The Δe shown in the illustration is equal to the total peak-to-peak ripple voltage. This is a complex wave which contains many harmonics of the fundamental ripple frequency and is the ripple that causes the noticeable heating of the capacitor.

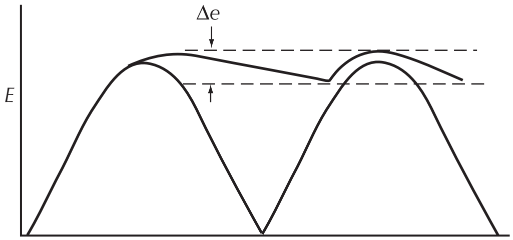


Figure 10-20. Capacitor charge and discharge on a full-wave rectifier output.

This can be mathematically determined or the ripple current through the capacitor can be measured by inserting a low impedance true rms ammeter in series with the capacitor. It is very important that the impedance of the meter be small compared with that of the capacitor, otherwise, a large measurement error will result.

- **Standard Life Tests.** *Standard life tests* at rated voltage and maximum rated temperatures are usually the criteria for determining the quality of an electrolytic capacitor. These two conditions rarely occur simultaneously in practice. Capacitor life expectancy is doubled for each 10°C (18°F) decrease in operating temperature, so a capacitor operating at room temperature will have a life expectancy 64 times that of the same capacitor operating at 85°C (185°F).
- **Surge Voltage.** The *surge voltage* specification of a capacitor determines its ability to withstand the high transient voltages that occur during the start up period of equipment. Standard tests specify a short on and long off period for an interval of 24 hours or more; the allowable surge voltage levels are usually 10% above the rated voltage of the capacitor. Fig. 10-21 shows how temperature, frequency, time, and applied voltage affect electrolytic capacitors.

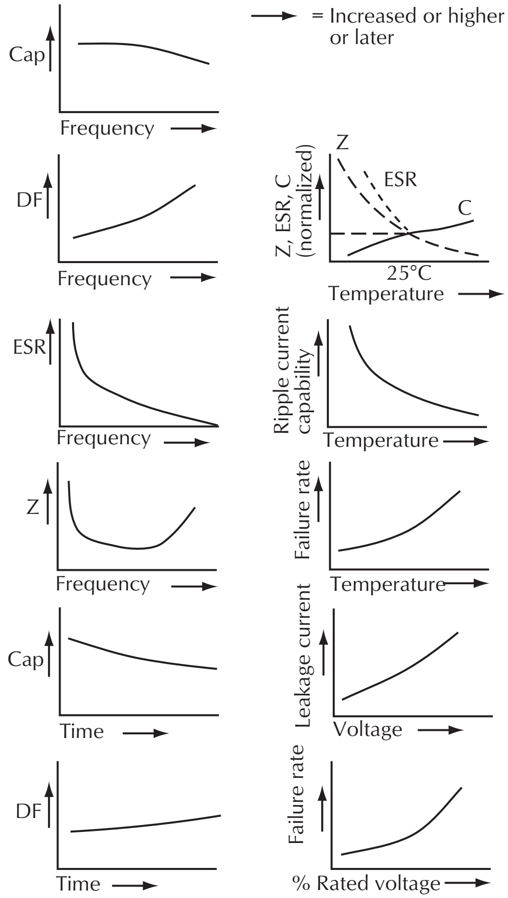


Figure 10-21. Variations in aluminum electrolytic characteristics caused by temperature, frequency, time, and applied voltage. Courtesy of Sprague Electric Company.

- **Tantalum Capacitors.** Tantalum electrolytics have become the preferred type where high reliability and long service life are paramount considerations.

Most metals form crystalline oxides that are nonprotecting, such as rust on iron or black oxide on copper. A few metals form dense, stable, tightly adhering, electrically insulating oxides. These are the so-called *valve* metals and include titanium, zirconium, niobium, tantalum, hafnium, and aluminum. Only a few of these permit the accurate control of oxide thickness by electrochemical means. Of these, the most valuable for the electronics industry are aluminum and tantalum.

The dielectric used in all tantalum electrolytic capacitors is tantalum pentoxide. Although wet foil capacitors use a porous paper separator between their foil plates, its function is merely to hold the electrolyte solution and to keep the foils from touching.

The tantalum pentoxide compound possesses high dielectric strength and a high dielectric constant. As capacitors are being manufactured, a film of tantalum pentoxide is applied to their electrodes by an electrolytic process. The film is applied in various thicknesses and at various voltages. Although transparent at first, it takes on different colors as light refracts through it. This coloring occurs on the tantalum electrodes of all three types of tantalum capacitors.

Rating for rating, tantalum capacitors tend to have as much as three times better capacitance/volume efficiency than aluminum electrolytic capacitors, because tantalum pentoxide has a dielectric constant of 26, some three times greater than that of aluminum oxide. This, in addition to the fact that extremely thin films can be deposited during manufacturing, makes the tantalum capacitor extremely efficient with respect to the number of microfarads available per unit volume.

The capacitance of any capacitor is determined by the surface area of the two conducting plates, the distance between the plates, and the dielectric constant of the insulating material between the plates.

The distance between the plates in tantalum electrolytic capacitors is very small since it is only the thickness of the tantalum pentoxide film. The dielectric constant of the tantalum pentoxide is high, therefore, the capacitance of a tantalum capacitor is high.

Tantalum capacitors contain either liquid or solid electrolytes. The liquid electrolyte in wet-slug and foil capacitors, usually sulfuric acid, forms the cathode or negative plate. In solid-electrolyte capacitors a dry material, manganese dioxide, forms the cathode plate.

The anode lead wire from the tantalum pellet consists of two pieces. A tantalum lead is embedded in, or welded to, the pellet, which is welded, in turn, to a nickel lead. In hermetically sealed types, the nickel lead is terminated to a tubular eyelet. An external lead of nickel or solder-coated nickel is soldered or welded to the eyelet. In encapsulated or plastic-encased designs, the nickel lead, which is welded to the basic tantalum lead, extends through the external epoxy resin coating or the epoxy end fill in the plastic outer shell.

Foil Tantalum Capacitors. *Foil tantalum capacitors* are made by rolling two strips of thin foil, separated by a paper saturated with electrolyte, into a convolute roll. The tantalum foil, which is to be the anode, is chemically etched to increase its effective surface area, providing more capacitance in a given volume. This is followed by anodizing in a chemical solution under direct voltage. This produces the dielectric tantalum pentoxide film on the foil surface.

Foil tantalum capacitors can be manufactured in dc working voltage values up to 300 V. However, of the three types of tantalum electrolytic capacitors, the foil design has the lowest capacitance per unit volume. It is also the least often encountered since it is best suited for the higher voltages primarily found in older designs of equipment and requires more manufacturing operations than do the two other types. Consequently, it is more expensive and is used only where neither a solid electrolyte nor a wet-slug tantalum capacitor can be employed.

Foil tantalum capacitors are generally designed for operation over the temperature range of −55°C to +125°C (−67°F to +257°F) and are found primarily in industrial and military electronics equipment.

Wet-Electrolyte Sintered Anode Tantalum Capacitors. *Wet-electrolyte sintered anode tantalum capacitors* often called *wet-slug* tantalum capacitors, use a pellet of sintered tantalum powder to which a lead has been attached. This anode has an enormous surface area for its size because of its construction. Tantalum powder of suitable fineness, sometimes mixed with binding agents, is machine-pressed into pellets. The second step is a sintering operation in which binders, impurities, and contaminants are vaporized and the tantalum particles are sintered into a porous mass with a very large internal surface area. A tantalum lead wire is attached by welding the wire to the pellet. (In some cases, the lead is embedded during pressing of the pellet before sintering.)

A film of tantalum pentoxide is electrochemically formed on the surface areas of the fused tantalum particles. The oxide is then grown to a thickness determined by the applied voltage.

Finally the pellet is inserted into a tantalum or silver container that contains an electrolyte solution. Most liquid electrolytes are gelled to prevent the free movement of the solution inside the container and to keep the electrolyte in intimate contact with the capacitor cathode. A suitable end seal arrangement prevents the loss of the electrolyte.

Wet-slug tantalum capacitors are manufactured in a working voltage range up to 150 Vdc.

Solid-Electrolyte Sintered Anode Tantalum Capacitors. *Solid-electrolyte sintered anode tantalum capacitors* differ from the wet versions in their electrolyte. Here, the electrolyte is manganese dioxide, which is formed on the tantalum pentoxide dielectric layer by impregnating the pellet with a solution of manganous nitrate. The pellets are then heated in an oven and the manganous nitrate is converted to manganese dioxide.

The pellet is next coated with graphite followed by a layer of metallic silver, which provides a solderable surface between the pellet and its can.

The pellets, with lead wire and header attached, are inserted into the can where the pellet is held in place by solder. The can cover is also soldered into place.

Another variation of the solid-electrolyte tantalum capacitor encases the element in plastic resins, such as epoxy materials. It offers excellent reliability and high stability for consumer and commercial electronics with the added feature of low cost.

Still other designs of solid tantalum capacitors, as they are commonly known, use plastic film or sleeving as the encasing material and others use metal shells which are back filled with an epoxy resin. And, of course, there are small tubular and rectangular molded plastic encasements as well.

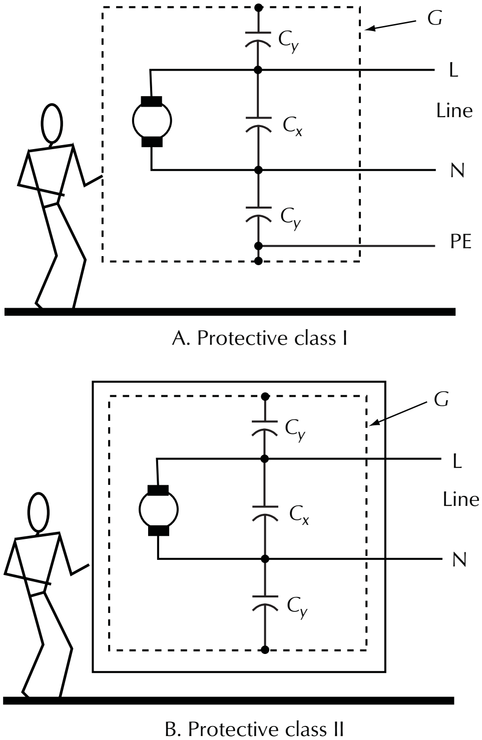
Tantalum Capacitors. In choosing between the three basic types of tantalum capacitors, the circuit designer customarily uses foil tantalum capacitors only where high voltage constructions are required or where there is substantial reverse voltage applied to a capacitor during circuit operation.

Wet-electrolyte sintered anode capacitors, or wet-slug tantalum capacitors, are used where the lowest dc leakage is required. The conventional silver can design will not tolerate any reverse voltages. However, in military or aerospace applications, tantalum cases are used instead of silver cases where utmost reliability is desired. The tantalum-cased wet-slug units will withstand reverse voltages up to 3 V, will operate under higher ripple currents, and can be used at temperatures up to 200°C (392°F).

Solid-electrolyte designs are the least expensive for a given rating and are used in many applications where their very small size for a given unit of capacitance is important. They will typically withstand up to 15% of the rated dc working voltage in a reverse direction. They also have good low temperature performance characteristics and freedom from corrosive electrolytes.

10.2.4.6 Suppression Capacitors

Suppression capacitors are used to reduce interference that comes in or out through the power line. They are effective because they are frequency dependent in that they become a short circuit at radio frequencies, without affecting low frequencies. Suppression capacitors are identified as X capacitors and Y capacitors. Fig. 10-22 shows two examples of radio interference suppression. Fig.10-22A is for protection class I which would include drills and hair dryers. Fig.10-22B is for protection class II where no protective conductor is connected to the metal case G.



A. Protective class I

B. Protective class II

Figure 10-22. Radio frequency suppression with X and Y capacitors. Courtesy of Vishay Roederstein.

X Capacitors. X capacitors are used across the mains to reduce symmetrical interference where a failure in the capacitor—i.e., the capacitor shorts out—will not cause injury, shock or death.

Y Capacitors. Y capacitors are used between a live conductor and a cabinet or case to reduce asymmetrical interference. Y capacitors have high electrical and mechanical specifications so they are much less likely to fail.

XY Capacitors. When used together they are called XY capacitors.

10.2.4.7 Supercapacitors

Supercapacitors, Ultracapacitors, more technically known as electrochemical double-layer capacitors, are one more step beyond the electrolytic capacitors. The charge-separation distance in ultracapacitors has been reduced to literally the dimensions of the ions within the electrolyte. In supercapacitors, the charges are not separated by millimeters or micrometers (microns) but by a few nanometers or from electrostatic capacitors to electrolytic capacitors to ultracapacitors. The charge-separation distance has in each instance dropped by three orders of magnitude, from 10^{-3} m to 10^{-6} m to 10^{-9} m.

- **How a Supercapacitor Works.** An supercapacitor or ultracapacitor, also known as a double-layer capacitor, polarizes an electrolytic solution to store energy electrostatically. Though it is an electrochemical device, no chemical reactions are involved in its energy storage mechanism. This mechanism is highly reversible and allows the ultracapacitor to be charged and discharged hundreds of thousands of times.

  An ultracapacitor can be viewed as two nonreactive porous plates, or collectors, suspended within an electrolyte, with a voltage potential applied across the collectors. In an individual ultracapacitor cell, the applied potential on the positive electrode attracts the negative ions in the electrolyte, while the potential on the negative electrode attracts the positive ions. A dielectric separator between the two electrodes prevents the charge from moving between the two electrodes.

  Once the ultracapacitor is charged and energy stored, a load can use this energy. The amount of energy stored is very large compared to a standard capacitor because of the enormous surface area created by the porous carbon electrodes and the small charge separation of 10 angstroms created by the dielectric separator. However, it stores a much smaller amount of energy than does a battery. Since the rates of charge and discharge are determined solely by its physical properties, the ultracapacitor can release energy much faster (with more power) than a battery that relies on slow chemical reactions.

  Many applications can benefit from ultracapacitors, whether they require short power pulses or low-power support of critical memory systems. Using an ultracapacitor in conjunction with a battery combines the power performance of the former with the greater energy storage capability of the latter. It can extend the life of a battery, save on replacement and maintenance costs, and enable a battery to be downsized. At the same time, it can increase available energy by providing high peak power whenever necessary. The combination of ultracapacitors and batteries requires additional dc/dc power electronics, which increases the cost of the circuit.

  Supercapacitors merged with batteries (hybrid battery) will become the new superbattery. Just about everything that is now powered by batteries will be improved by this much better energy supply. They can be made in most any size, from postage stamp to hybrid car battery pack. Their light weight and low cost make them attractive for most portable electronics and phones, as well as for aircraft and automobiles.

- **Advantages of a Supercapacitor**

  1. Virtually unlimited life cycle—cycles millions of times—10 to 12 year life.
  2. Low internal impedance.
  3. Can be charged in seconds.
  4. Cannot be overcharged.
  5. Capable of very high rates of charge and discharge.
  6. High cycle efficiency (95% or more).

- **Disadvantages of a Supercapacitor:**

  1. Supercapacitors and ultra capacitors are relatively expensive in terms of cost per watt.
  2. Linear discharge voltage prevents use of the full energy spectrum.
  3. Low energy density—typically holds one-fifth to one-tenth the energy of an electrochemical battery.
  4. Cells have low voltages; therefore, serial connections are needed to obtain higher voltages, which require voltage balancing if more than three capacitors are connected in series.
  5. High self-discharge—the self-discharge rate is considerably higher than that of an electrochemical battery.
  6. Requires sophisticated electronic control and switching equipment.

A supercapacitor by itself cannot totally replace the battery. But, by merging a supercapacitor and a battery together—like a hybrid battery, it will be possible for supercapacitors to replace the battery as we know it today.

Presently supercapacitors need batteries to store the energy and are basically used as a buffer between the battery and the device. Supercapacitors can be charged and discharged hundreds of thousands of times where a battery cannot do that.

- **Calculating Backup Time.** To calculate the desired backup time the supercapacitor will provide if the power goes off, the starting and ending voltage on the capacitor, the current draw from the capacitor, and the capacitor size must be known.

  Assuming that the load draws a constant current while running from V_{BACKUP}, then the worst-case backup time in hours would use the equation:

$$Backup\ time = \frac{\dfrac{C(V_{BACKUPSTART} - V_{BACKUPMIN})}{I_{BACKUPMAX}}}{3600} \tag{10-24}$$

where,

C is the capacitor value in farads,

$V_{BACKUPSTART}$ is the initial voltage in volts. The voltage applied to V_{CC}, less the voltage drop from the diodes, if any, used in the charging circuit,

$V_{BACKUPMIN}$ is the ending voltage in volts,

$I_{BACKUPMAX}$ is the maximum V_{BACKUP} current in amperes.

For example, to determine how long the backup time will be under the following conditions:

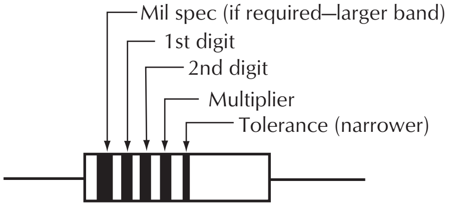
- 0.2 F capacitor
- $V_{BACKUPSTART}$ is 3.3 V
- $V_{BACKUPMIN}$ is 1.3 V
- $I_{BACKUPMAX}$ is 1000 nA, then:

$$Backup\ time = \frac{\dfrac{0.2(3.3 - 1.3)}{10^{-6}}}{300}$$

$$= 111.1\ \text{h}$$

10.3 Inductors

Inductance is used for the storage of electrical energy in a magnetic field, called *magnetic energy*. Magnetic energy is stored as long as current keeps flowing through the inductor. The current of a sine wave lags the voltage by 90° in a perfect inductor. Figure 10-23 shows the color code for small inductors.



Color	Inductance digits 1st	2nd	Multiplier	Tolerance
Black	0	0	1	
Brown	1	1	10	
Red	2	2	100	
Orange	3	3	1000	
Yellow	4	4	10,000	
Green	5	5		
Blue	6	6		
Violet	7	7		
Gray	8	8		
White	9	9		
Gold	–	–		±5%
Silver	–	–	–	±10%
No band	–	–	–	±20%

Figure 10-23. Color code for small inductors (in μH).

10.3.1 *Types of Inductors*

Inductors are constructed in a variety of ways, depending on their use.

10.3.1.1 Air Core Inductors

Air core inductors are either ceramic core or phenolic core.

10.3.1.2 Axial Inductor

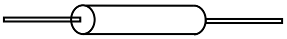
An *axial inductor* is constructed on a core with concentric leads on opposite ends, Fig.10-24A. The core material may be phenolic, ferrite, or powdered iron.

10.3.1.3 Bobbin Core Inductor

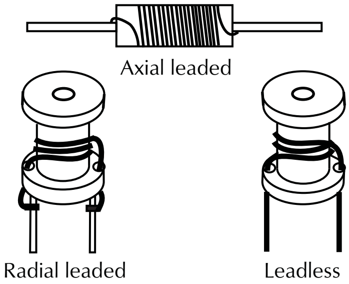
Bobbin core inductors have the shape of a bobbin and may come with or without leads. They may be either axial or radial, Fig. 10-24B.

10.3.1.4 Ceramic Core

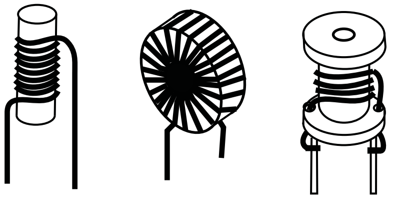
Ceramic core inductors are often used in high frequency applications where low inductance, low core losses, and high Q values are required. Ceramic has no magnetic properties so there is no increase in permeability due to the core material.



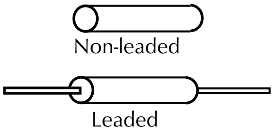
A. Axial leaded inductor.



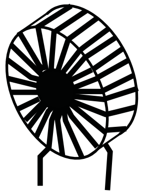
B. Bobbins.



C. Radial inductors.



D. Slug cores.



E. Toroidal core.

Figure 10-24. Various inductor core types.

Ceramic has a low thermal coefficient of expansion allowing high inductance stability over a high operating temperature range.

10.3.1.5 Epoxy-Coated Inductor

Epoxy-coated inductors usually have a smooth surface and edges. The coating provides insulation.

10.3.1.6 Ferrite Core

Ferrite cores can be easily magnetized. The core consists of a mixture of oxide of iron and other elements such as manganese and zinc (MnZn) or nickel and zinc (NiZn). The general composition is $xxFe_2O_4$ where xx is one of the other elements.

10.3.1.7 Laminated Cores

Laminated cores are made by stacking insulated laminations on top of each other. Some laminations have the grains oriented to minimize core losses, giving higher permeability. Laminated cores are more common in transformers.

10.3.1.8 Molded Inductor

A *molded inductor* has its case formed via a molding process, creating a smooth, well-defined body with sharp edges.

10.3.1.9 MPP Core

MPP, or moly perm alloy powder, is a magnetic material with a inherent distributed air gap, allowing it to store higher levels of magnetic flux compared to other materials. This allows more dc to flow through the inductor before the core saturates.

The core consists of 80% nickel, 2–3% molybdenum, and the remaining percentage iron.

10.3.1.10 Multilayer Inductor

A *multilayer inductor* consists of layers of coil between layers of core material. The coil is usually bare metal and is sometimes referred to as *nonwirewound*.

10.3.1.11 Phenolic Core

Phenolic cores are often called *air cores* and are often used in high frequency applications where low inductance values, low core losses, and high Q values are required.

Phenolic has no magnetic properties so there is no increase in permeability due to the core material. Phenolic cores provide high strength, high flammability ratings, and high temperature characteristics.

10.3.1.12 Powdered Iron Core

Powdered iron is a magnetic material with an inherent distributed air gap that allows the core to have high levels of magnetic flux. This allows a high level of dc to flow through the core before saturation.

Powdered iron cores are close to 100% iron whose particles are insulated and mixed with a binder such as epoxy or phenolic. They are pressed into a final shape and cured by baking.

10.3.1.13 Radial Inductor

A *radial inductor* is constructed on a core with leads on the same side, Fig. 10-24C. This allows for easy mounting on circuit boards, etc.

10.3.1.14 Shielded Inductor

A *shielded inductor* has its core designed to contain the majority of the magnetic field. Some are self shielding such as toroids, e-cores, and pot cores. Bobbin and slug cores require a magnetic sleeve for shielding.

10.3.1.15 Slug Core

Slug cores have the shape of a cylindrical rod and come with or without leads, Fig.10-24D. They have higher flux density characteristics than other core shapes as most of the magnetic energy is stored in the air around the core.

10.3.1.16 Tape Wound Core

Tape wound cores are made by rolling insulated and precisely controlled thickness strips of alloy iron into a toroidal shape. The finished cores have an outside coating for protection.

Tape wound cores are capable of storing high amounts of energy and contain a high permeability.

10.3.1.17 Toroidal Inductor

Toroidals are constructed by placing the winding on a donut-shaped core, Fig. 10-24E. Toroidal cores may be ferrite, powdered iron, tape wound, or alloy and high flux.

Toroidals are self shielding andhave efficient energy transfer, high coupling between windings, and early saturation.

10.3.2 Impedance Characteristics

Impedance. The impedance or inductive reactance (X_L) of an inductor to an ac signal is found with the equation

$$X_L = 2\pi fL \tag{10-25}$$

where,
f is the frequency in hertz,
L is the inductance in henrys.

The inductance of a coil is only slightly affected by the type of wire used for its construction. The Q of the coil will be governed by the ohmic resistance of the wire. Coils wound with silver or gold wire have the highest Q for a given design.

To increase the inductance, inductors can be connected in series. The total inductance will always be greater than the largest inductor

$$L_T = L_1 + L_2 \cdots + L_n \tag{10-26}$$

To reduce the total inductance, place the inductors in parallel. The total inductance will always be less than the value of the lowest inductor

$$L_T = \frac{1}{\frac{1}{L_1} + \frac{1}{L_2} \cdots + \frac{1}{L_n}} \tag{10-27}$$

To determine the impedance of circuits with resistance, capacitance, and inductance, see Section 10.4.

Mutual Inductance. *Mutual inductance* is the property that exists between two conductors that are carrying current when the magnetic lines of force from one conductor link with the magnetic lines of force of the other. The mutual inductance of two coils with fields interacting can be determined by

$$M = \frac{L_A - L_B}{4} \tag{10-28}$$

where,
M is the mutual inductance of L_A and L_B in henrys,
L_A is the total inductance of coils L_1 and L_2 with fields aiding in henrys,
L_B is the total inductance of coils L_1 and L_2 with fields opposing in henrys.

The coupled inductance can be determined by the following equations.

In parallel with fields aiding

$$L_T = \frac{1}{\frac{1}{L_1 + M} + \frac{M}{L_2 + M}} \tag{10-29}$$

In parallel with fields opposing

$$L_T = \frac{1}{\frac{1}{L_1 - M} - \frac{M}{L_2 - M}} \tag{10-30}$$

In series with fields aiding

$$L_T = L_1 + L_2 + 2M \tag{10-31}$$

In series with fields opposing

$$L_T = L_1 + L_2 - 2M \tag{10-32}$$

where,
L_T is the total inductance in henrys,
L_1 and L_2 are the inductances of the individual coils in henrys,
M is the mutual inductance in henrys.

When two coils are inductively coupled to give transformer action, the coupling coefficient is determined by

$$K = \frac{M}{\sqrt{L_1 \times L_2}} \tag{10-33}$$

where,
K is the coupling coefficient,
M is the mutual inductance in henrys,
L_1 and L_2 are the inductances of the two coils in henrys.

An inductor in a circuit has a reactance of $j2\pi fL\ \Omega$. Mutual inductance in a circuit also has a reactance equal to $j2\pi fM\ \Omega$. The operator j denotes reactance. The energy stored in an inductor can be determined by

$$W = \frac{LI^2}{2} \tag{10-34}$$

where,
W is the energy in joules (watt-seconds),
L is the inductance in henrys,
I is the current in amperes.

Coil Inductance. The following is the relationship of the turns in a coil to its inductance:

- The inductance is proportional to the square of the turns.
- The inductance increases as the permeability of the core material is increased.
- The inductance increases as the cross-sectional area of the core material is increased.
- The inductance increases as the length of the winding is increased.
- A shorted turn decreases the inductance. In an audio transformer, the frequency characteristic will be affected, and the insertion loss increased.
- Inserting an iron core in a coil increases the inductance; hence, its inductive reactance is increased.
- Introducing an air gap in an iron core coil reduces the inductance.

The maximum voltage induced in a conductor moving in a magnetic field is proportional to the number of magnetic lines of force cut by the conductor moving in the field. A conductor moving parallel to the lines of force cuts no lines of force so no current is generated in the conductor. A conductor moving at right angles to the lines of force will cut the maximum number of lines per inch per second; therefore, the voltage will be at the maximum.

A conductor moving at any angle to the lines of force cuts a number of lines of force proportional to the sine of the angles.

$$V = \beta L v \sin\theta \times 10^{-8} \tag{10-35}$$

where,
V is the voltage produced,
β is the flux density,
L is the length of the conductor in centimeters,
v is the velocity in centimeters per second of the conductor moving at an angle θ.

The direction of the induced electromotive force (*emf*) is in the direction in which the axis of a right-hand screw, when turned with the velocity vector, moves through the smallest angle toward the flux density vector. This is called the *right-hand rule*.

The magnetomotive force produced by a coil is derived by

$$\begin{aligned} \textit{ampere turns} &= T\left(\frac{V}{R}\right) \\ &= TI \end{aligned} \tag{10-36}$$

where,
T is the number of turns,
V is the voltage in volts,
R is the resistance of the wire in ohms,

I is the current in amperes.

The inductance of single-layer, spiral, and multilayer coils can be calculated by using either Wheeler's or Nagaoka's equations. The accuracy of the calculation will vary between 1% and 5%. The inductance of a single-layer coil, Fig. 10-25A, may be found using Wheeler's equation

$$L = \frac{B^2N^2}{9B + 10A} \tag{10-37}$$

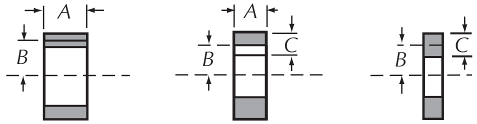
For the multilayer coil, Fig. 10-25B, the calculations are

$$L = \frac{0.8B^2N^2}{6B + 9A + 10C} \tag{10-38}$$

For the spiral coil, Fig. 10-25C, the calculations are:

$$L = \frac{B^2N^2}{8B + 11C} \tag{10-39}$$

where,
B is the radius of the winding,
N is the number of turns in the coil,
A is the length of the winding,
C is the thickness of the winding,
L is in μH.



A. Single layer B. Multilayer C. Spiral

Figure 10-25. Single- and multilayer inductors.

Q. Q is the ratio of the inductive reactance to the internal resistance of the coil. The principal factors that affect Q are frequency, inductance, dc resistance, inductive reactance, and the type of winding. Other factors are the core losses, the distributed capacity, and the permeability of the core material. The Q for a coil where R and L are in series is

$$Q = \frac{2\pi fL}{R} \tag{10-40}$$

where,
f is the frequency in hertz,
L is the inductance in henrys,
R is the resistance in ohms.

The Q of the coil can be measured as follows. Using the circuit of Fig. 10-26, Q of a coil may be easily measured for frequencies up to 1 MHz. Since the voltage across an inductance at resonance equals $Q \times V$, where V is the voltage developed by the oscillator, it is necessary only to measure the output voltage from the oscillator and the voltage across the inductance.

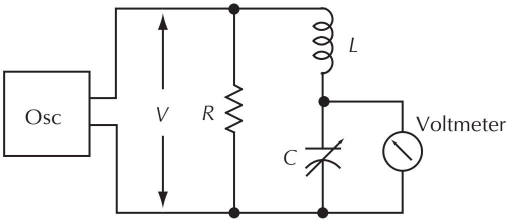


Figure 10-26. Circuit for measuring the Q of a coil.

The voltage from the oscillator is introduced across a low value of resistance R, about 1% of the anticipated radiofrequency resistance of the LC combination, to assure that the measurement will not be in error by more than 1%. For average measurements, resistor R will be on the order of 0.10 Ω. If the oscillator cannot be operated into an impedance of 0.10 Ω, a matching transformer may be employed. It is desirable to make C as large as convenient to minimize the ratio of the impedance looking from the voltmeter to the impedance of the test circuit. The voltage across R is made small, on the order of 0.10 V. The LC circuit is then adjusted to resonate and the resultant voltage measured. The value of Q may then be equated

$$Q = \frac{\textit{Resonant voltage across C}}{\textit{Voltage across R}}. \tag{10-41}$$

The Q of a coil may be approximated by the equation

$$Q = \frac{2\pi fL}{R} = \frac{X_L}{R} \tag{10-42}$$

where,
f is the frequency in hertz,
L is the inductance in henrys,
R is the dc resistance in ohms as measured by an ohmmeter,
X_L is the inductive reactance of the coil.

Time Constant. When a dc voltage is applied to an RL circuit, a certain amount of time is required to change the voltage. In a circuit containing inductance and resistance, the time constant is defined as the time it takes for the current to reach 62.3% of its maximum value. The time constant can be determined with the following equation:

$$T = \frac{L}{R} \tag{10-43}$$

where,

T is the time in seconds,

L is the inductance in henrys,

R is the resistance in ohms.

See Section 10.2.1 for a further discussion of time constants. The effect of an inductor is the same as for a capacitor and resistor. Also, curve A in Fig. 10-14 shows the current through an inductor on buildup and curve B shows the current decay when the voltage is removed.

Right-Hand Rule. The *right-hand rule* is a method devised for determining the direction of a magnetic field around a conductor carrying a direct current. The conductor is grasped in the right hand with the thumb extended along the conductor. The thumb points in the direction of the current. If the fingers are partly closed, the fingertips will point in the direction of the magnetic field.

Maxwell's rule states, "If the direction of travel of a right-handed corkscrew represents the direction of the current in a straight conductor, the direction of rotation of the corkscrew will represent the direction of the magnetic lines of force."

10.3.3 Ferrite Beads

The original ferrite beads were small round ferrites with a hole through the middle where a wire passed through. Today they come as the original style plus as multiple apertures and surface mount configurations.

The ferrite bead can be considered a frequency-dependent resistor whose equivalent circuit is a resistor in series with an inductor. As the frequency increases, the inductive reactance increases and then decreases, and the complex impedance of the ferrite material increases the overall impedance of the bead, Fig. 10-27.

At frequencies below 10 MHz, the impedance is less than 10 Ω. As the frequency increases, the impedance increases to about 100 Ω and becomes mostly resistive at 100 MHz.

Once the impedance is resistive, resonance does not occur as it would using an LC network. Ferrite beads do not attenuate low frequencies or dc so are useful for reducing EMI/EMC in audio circuits.

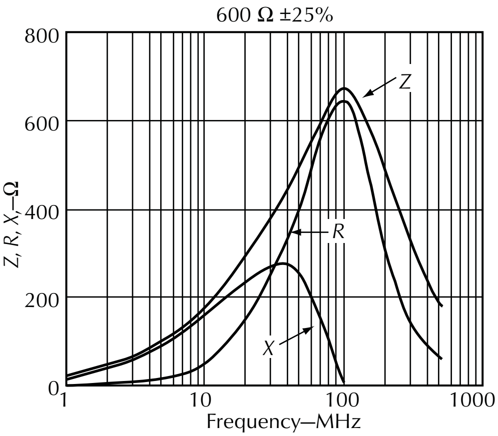


Figure 10-27. Impedance of ferrite beads. Courtesy of Vishay Dale.

10.3.4 Skin Effect

Skin effect is the tendency of ac to flow near the surface of a conductor rather than flowing through the conductor's entire cross sectional area. This increases the resistance of the conductor because the magnetic field caused by the current creates eddy currents near the center of the conductor. The eddy currents oppose the normal flow of current near the center, forcing the main current flow out toward the surface as the frequency of the ac current increases.

To reduce this problem, a wire made up of separately insulated strands woven and/or bunched together is used. Commonly called Litz wire, the current is equally divided between all of the individual strands which equalizes the flux linkage and reactance of the individual strands, reducing the ac losses compared to solid wire.

10.3.5 Shielded Inductor

Some inductor designs are self-shielding. Examples are toroid, pot core, and E-core inductors. Slug cores and bobbins may require shielding, depending on the application. It is impossible to completely shield an inductor.

10.4 Impedance

The total impedance created by resistors, capacitors, and inductors in circuits can be determined with the following equations.

Parallel circuits

$$Z = \frac{RX}{\sqrt{R^2 + X^2}} \tag{10-44}$$

Series circuits

$$Z = \sqrt{R^2 + X^2} \tag{10-45}$$

Resistance and inductance in series

$$Z = \sqrt{R^2 + {X_L}^2} \tag{10-46}$$

$$\theta = \operatorname{atan}\frac{X_L}{R} \tag{10-47}$$

Resistance and capacitance in series

$$Z = \sqrt{R^2 + {X_C}^2} \tag{10-48}$$

$$\theta = \operatorname{atan}\frac{X_C}{R} \tag{10-49}$$

Inductance and capacitance in series when X_L is larger than X_C

$$Z = X_L - X_C \tag{10-50}$$

Inductance and capacitance in series when X_C is larger than X_L

$$Z = X_C - X_L \tag{10-51}$$

Resistance, inductance, and capacitance in series

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \tag{10-52}$$

$$\theta = \operatorname{atan}\frac{X_L - X_C}{R} \tag{10-53}$$

Resistance and inductance in parallel

$$Z = \frac{RX_L}{\sqrt{R^2 + {X_L}^2}} \tag{10-54}$$

Capacitance and resistance in parallel

$$Z = \frac{RX_C}{\sqrt{R^2 + {X_C}^2}} \tag{10-55}$$

Capacitance and inductance in parallel when X_L is larger than X_C

$$Z = \frac{X_L \times X_C}{X_L - X_C} \tag{10-56}$$

Capacitance and inductance in parallel when X_C is larger than X_L

$$Z = \frac{X_C \times X_L}{X_C - X_L} \tag{10-57}$$

Inductance, capacitance, and resistance in parallel

$$Z = \frac{RX_LX_C}{\sqrt{{X_L}^2{X_C}^2 + R^2(X_L - X_C)^2}} \tag{10-58}$$

$$\theta = \operatorname{atan}\frac{R(X_L - X_C)}{X_LX_C} \tag{10-59}$$

Inductance and series resistance in parallel with resistance

$$Z = R_2\sqrt{\frac{{R_1}^2 + {X_L}^2}{(R_1 + R_2)^2 + {X_L}^2}} \tag{10-60}$$

$$\theta = \operatorname{atan}\frac{R_2X_L}{{R_1}^2 + {X_L}^2 + R_1R_2} \tag{10-61}$$

Inductance and series resistance in parallel with capacitance

$$Z = X_C\sqrt{\frac{R^2 + {X_L}^2}{R^2 + (X_L - X_C)^2}} \tag{10-62}$$

$$\theta = \operatorname{atan}\frac{X_L(X_C - X_L) - R^2}{RX_C} \tag{10-63}$$

Capacitance and series resistance in parallel with inductance and series resistance

$$Z = \sqrt{\frac{({R_1}^2 + {X_L}^2)({R_2}^2 + {X_C}^2)}{(R_1 + R_2)^2 + (X_L - X_C)^2}} \qquad (10\text{-}64)$$

$$Z = \operatorname{atan}\frac{X_L({R_2}^2 + {X_C}^2) - X_C({R_1}^2 + {X_L}^2)}{R_1({R_2}^2 + {X_C}^2) + R_2({R_1}^2 + {X_L}^2)} \qquad (10\text{-}65)$$

where,

Z is the impedance in ohms,

R is the resistance in ohms,

L is the inductance in henrys,

X_L is the inductive reactance in ohms,

X_C is the capacitive reactance in ohms.

θ is the phase angle in degrees by which the current leads the voltage in a capacitive circuit or lags the voltage in an inductive circuit. 0° indicates an in-phase condition.

10.5 Resonant Frequency

When an inductor and capacitor are connected in series or parallel, they form a resonant circuit. The resonant frequency can be determined from the equation

$$\begin{aligned} f &= \frac{1}{2\pi\sqrt{LC}} \\ &= \frac{1}{2\pi C X_C} \qquad (10\text{-}66) \\ &= \frac{X_L}{2\pi L} \end{aligned}$$

where,
L is the inductance in henrys,
C is the capacitance in farads,
X_L and X_C are the impedance in ohms.

The resonant frequency can also be determined through the use of a reactance chart developed by the Bell Telephone Laboratories, Fig. 10-28. This chart can be used for solving problems of inductance, capacitance, frequency, and impedance. If two of the values are known, the third and fourth values may be found with its use. As an example, what is the value of capacitance and inductance required to resonate at a frequency of 1000 Hz in a circuit having an impedance of 500 Ω? Entering the chart on the 1000 Hz vertical line and following it to the 500 Ω line (impedance is shown along the left-hand margin), the value of inductance is indicated by the diagonal line running upward as 0.08 H (80 mH), and the capacitance indicated by the diagonal line running downward at the right-hand margin is 0.3 µF.

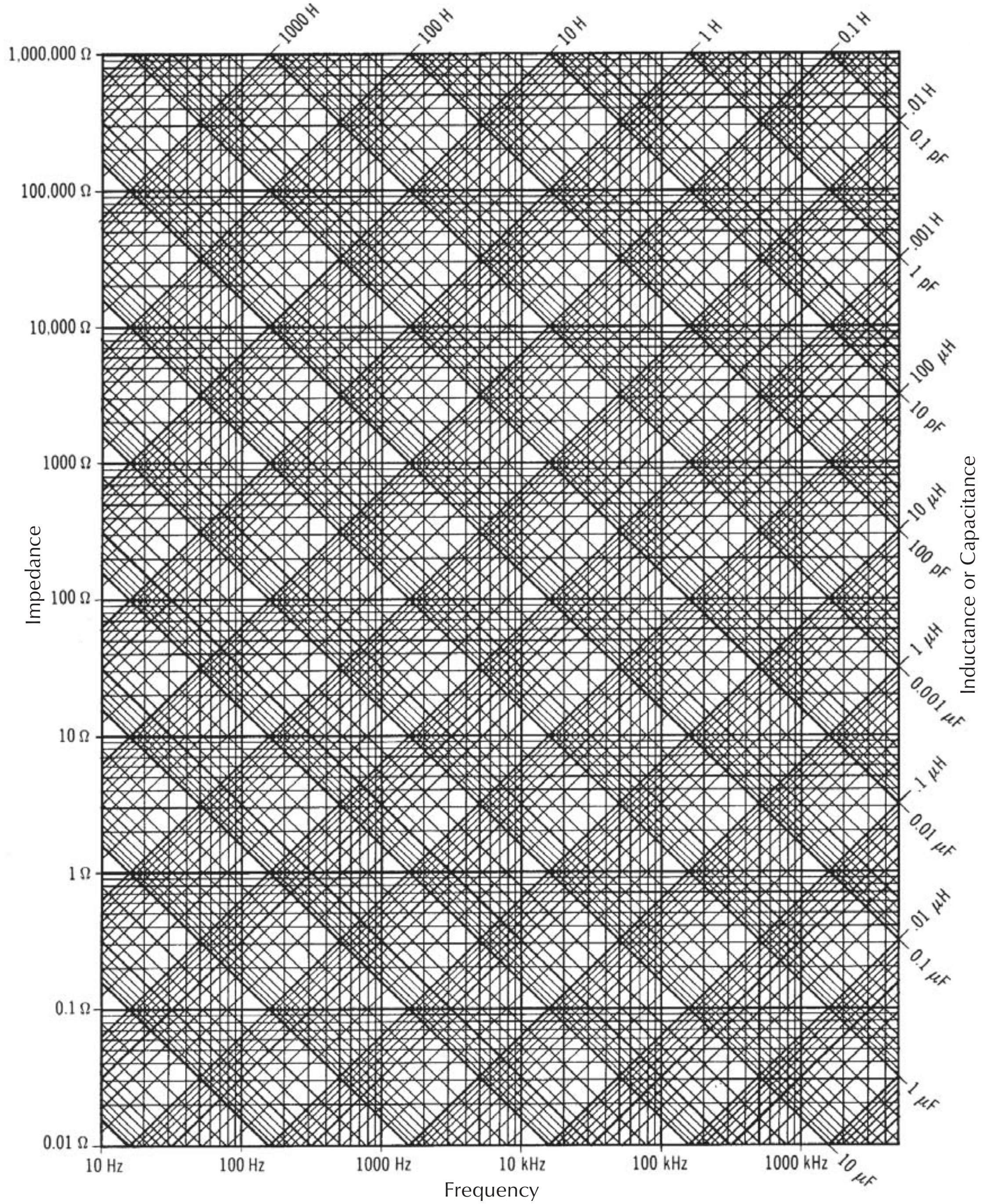


Figure 10-28. Reactance chart. Courtesy AT&T Bell Laboratories.

References

Inductor and Magnetic Product Terminology, Document No. 34053, March 2002, Vishay Dale.

ILB, ILBB Ferrite Beads, Document No. 34091, May 1999, Vishay Dale.

Ohmite Manual of Engineering Information, Bulletin 1100, Ohmite Mfg. Co.

Ohmite Resistor Selection Design Data, Ohmite Mfg. Co.

Radio Interference Suppression Capacitors, Document No. 26529, September 2002, Vishay Roederstein.

Chapter 11

Audio Transformer Basics

by Bill Whitlock

11.1 Audio Transformer Basics . . . 275
  11.1.1 Basic Principles and Terminology . . . 275
    11.1.1.1 Magnetic Fields and Induction . . . 275
    11.1.1.2 Windings and Turns Ratio . . . 276
    11.1.1.3 Excitation Current . . . 276
  11.1.2 Realities of Practical Transformers . . . 277
    11.1.2.1 Core Materials and Construction . . . 278
    11.1.2.2 Winding Resistances and Auto-Transformers . . . 280
    11.1.2.3 Leakage Inductance and Winding Techniques . . . 281
    11.1.2.4 Winding Capacitances and Faraday Shields . . . 282
    11.1.2.5 Magnetic Shielding . . . 283
  11.1.3 General Application Considerations . . . 283
    11.1.3.1 Maximum Signal Level, Distortion, and Source Impedance . . . 283
    11.1.3.2 Frequency Response . . . 285
    11.1.3.3 Insertion Loss . . . 286
    11.1.3.4 Sources with Zero Impedance . . . 287
    11.1.3.5 Bi-Directional Reflection of Impedances . . . 287
    11.1.3.6 Transformer Noise Figure . . . 288
    11.1.3.7 Basic Classification by Application . . . 289
11.2 Audio Transformers for Specific Applications . . . 290
  11.2.1 Equipment-Level Applications . . . 290
    11.2.1.1 Microphone Input . . . 290
    11.2.1.2 Line Input . . . 291
    11.2.1.3 Moving-Coil Phono Input . . . 291
    11.2.1.4 Line Output . . . 292
    11.2.1.5 Inter-Stage and Power Output . . . 293
    11.2.1.6 Microphone Output . . . 294
  11.2.2 System-Level Applications . . . 295
    11.2.2.1 Microphone Isolation or Splitter . . . 295
    11.2.2.2 Microphone Impedance Conversion . . . 295
    11.2.2.3 Line to Microphone Input or Direct Box . . . 296
    11.2.2.4 Line Isolation or Hum Eliminators . . . 296
    11.2.2.5 Loudspeaker Distribution or Constant Voltage . . . 298
    11.2.2.6 Telephone Isolation or Repeat Coil . . . 300
    11.2.2.7 Telephone Directional Coupling or Hybrid . . . 300
    11.2.2.8 Moving-Coil Phono Step-Up . . . 301
11.3 Measurements and Data Sheets . . . 302
  11.3.1 Testing and Measurements . . . 302
    11.3.1.1 Transmission Characteristics . . . 302

11.3.1.2 Balance Characteristics 302
11.3.1.3 Resistances, Capacitances, and Other Data 303
11.3.2 Data Sheets 303
11.3.2.1 Data to Impress or to Inform? 303
11.3.2.2 Comprehensive Data Sheet Example 303
11.4 Installation and Maintenance 306
11.4.1 A Few Installation Tips. 306
11.4.2 De-Magnetization 306
References 307

11.1 Audio Transformer Basics

Since the birth of audio electronics, the audio transformer has played an important role. When compared to modern miniaturized electronics, a transformer seems large, heavy, and expensive but it continues to be the most effective solution in many audio applications. The usefulness of a transformer lies in the fact that electrical energy can be transferred from one circuit to another without direct connection (e.g., isolation from ground loops), and in the process the energy can be readily changed from one voltage level to another (e.g., impedance matching). Although a transformer is not a complex device, considerable explanation is required to properly understand how it operates. This chapter is intended to help the audio system engineer properly select and apply transformers. In the interest of simplicity, only basic concepts of their design and manufacture will be discussed.

11.1.1 Basic Principles and Terminology

11.1.1.1 Magnetic Fields and Induction

As shown in Fig. 11-1, a *magnetic field* is created around any conductor (wire) in which current flows. The strength of the field is directly proportional to current. These invisible magnetic *lines of force*, collectively called *flux*, are set up at right angles to the wire and have a direction, or magnetic polarity, that depends on the direction of current flow. Note that although the flux around the upper and lower wires have different directions, the lines inside the loop aid because they point in the same direction. If an alternating current flows in the loop, the instantaneous intensity and polarity of the flux will vary at the same frequency and in direct proportion to Fig. 11-2, as expanding, contracting, and reversing in polarity with each cycle of the ac current. The law of induction states that a voltage will be *induced* in a conductor exposed to changing flux and that the induced voltage will be proportional to the rate of the flux change. This voltage has an instantaneous polarity which opposes the original current flow in the wire, creating an apparent resistance called *inductive reactance*. Inductive reactance is calculated according to the formula

$$X_L = 2\pi fL \tag{11-1}$$

where,

X_L is inductive reactance in ohms,

f is the frequency in hertz,

L is inductance in Henrys.

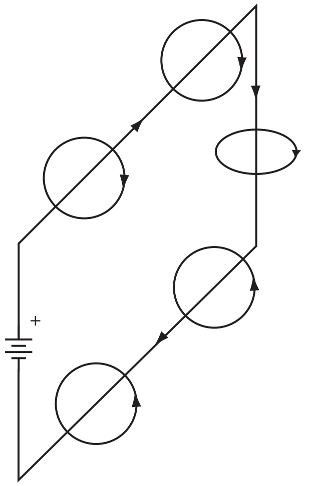


Figure 11-1. Magnetic field surrounding conductor.

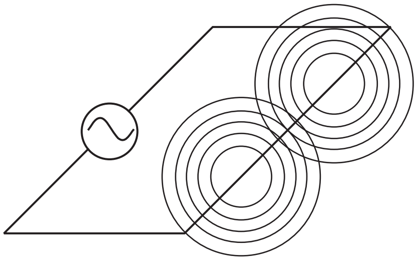


Figure 11-2. ac magnetic field.

An *inductor* generally consists of many turns or loops of wire called a *coil*, as shown in Fig. 11-3, which links and concentrates magnetic flux lines, increasing the *flux density*. The inductance of any given coil is determined by factors such as the number of turns, the physical dimensions and nature of the winding, and the properties of materials in the path of the magnetic flux.

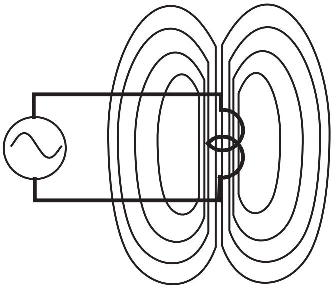


Figure 11-3. Coil concentrates flux.

According to the law of induction, a voltage will be induced in any conductor (wire) that cuts flux lines.

Therefore, if we place two coils near each other as shown in Fig. 11-4, an ac *current* in one coil will induce an ac voltage in the second coil. This is the essential principle of energy transfer in a *transformer*. Because they require a changing magnetic field to operate, transformers will not work at dc. In an ideal transformer, the magnetic coupling between the two coils is total and complete, i.e., all the flux lines generated by one coil cut across all the turns of the other. The *coupling coefficient* is said to be unity or 1.00.

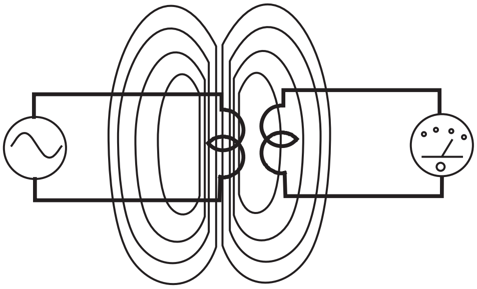


Figure 11-4. Inductive coupling.

11.1.1.2 Windings and Turns Ratio

The coil or winding that is driven by an electrical source is called the *primary* and the other is called the *secondary*. The ratio of the number of turns on the primary to the number of turns on the secondary is called the *turns ratio*. Since essentially the same voltage is induced in each turn of each winding, the primary to secondary *voltage ratio* is the same as the turns ratio. For example, with 100 turns on the primary and 50 turns on the secondary, the turns ratio is 2:1. Therefore, if 20 V were applied to the primary, 10 V would appear at the secondary. Since it reduces voltage, this transformer would be called a *step-down* transformer. Conversely, a transformer with a turns ratio of 1:2 would be called a *step-up* transformer since its secondary voltage would be twice that of the primary. Since a transformer cannot create power, the power output from the secondary of an ideal transformer can only equal (and in a real transformer can only be less than) the power input to the primary. Consider an ideal 1:2 step-up transformer. When 10 V is applied to its primary, 20 V appears at its secondary. Since no current is drawn by the primary (this is an ideal transformer—see "11.1.1.3, Excitation Current," its impedance appears to be infinite or an open circuit.

However, when a 20 Ω load is connected to the secondary, a current of 1 A flows making output power equal 20 W. To do this, a current of 2 A must be drawn by the primary, making input power equal 20 W. Since the primary is now drawing 2 A with 10 V applied, its impedance appears to be 5 Ω. In other words, the 20 Ω load impedance on the secondary has been *reflected* to the primary as 5 Ω. In this example, a transformer with a 1:2 turns ratio exhibited an impedance ratio of 1:4. Transformers always reflect impedances from one winding to another by the square of their turns ratio or, expressed as an equation

$$\frac{Z_p}{Z_s} = \left(\frac{N_p}{N_s}\right)^2 \tag{11-2}$$

where,

Z_p is primary impedance,

Z_s is secondary impedance,

N_p/N_s is turns ratio, which is the same as the voltage ratio.

When a transformer converts voltage, it also converts impedance—and vice versa.

The direction in which coils are wound—i.e., clockwise or counterclockwise—and/or the connections to the start or finish of each winding determines the instantaneous *polarity* of the ac voltages. All windings that are wound in the same direction will have the same polarity between start and finish ends. Therefore, relative to the primary, polarity can be inverted by either (1) winding the primary and secondary in opposite directions, or (2) reversing the start and finish connections to either winding. In schematic symbols for transformers, dots are generally used to indicate which ends of windings have the same polarity. Observing polarity is essential when making series or parallel connections to transformers with multiple windings. *Taps* are connections made at any intermediate point in a winding. For example, if 50 turns are wound, an electrical connection brought out, and another 50 turns completes the winding, the 100 turn winding is said to be *centertapped*.

11.1.1.3 Excitation Current

While an ideal transformer has infinite primary inductance, a real transformer does not. Therefore, as shown in Fig. 11-5, when there is no load on the secondary and an ac voltage is applied to the primary, an *excitation current* will flow in the primary, creating magnetic excitation flux around the winding. In theory, the current is due only to the inductive reactance of the primary winding. In accordance with Ohm's Law and the equation for inductive reactance,

$$I_E = \frac{E_p}{2\pi f L_p} \tag{11-3}$$

where,

I_E is excitation current in amperes,

E_P is primary voltage in volts,

f is frequency in hertz,

L_P is primary inductance in henrys.

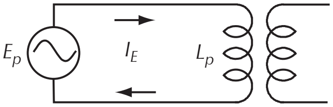


Figure 11-5. Excitation current.

Obviously, if primary inductance were infinite, excitation current would be zero. As shown in Fig. 11-6, when a load is connected, current will flow in the secondary winding. Because secondary current flows in the opposite direction, it creates magnetic flux which opposes the excitation flux. This causes the impedance of the primary winding to drop, resulting in additional current being drawn from the driving source. Equilibrium is reached when the additional flux is just sufficient to completely cancel that created by the secondary. The result, which may surprise some, is that flux density in a transformer is *not* increased by load current. This also illustrates how load current on the secondary is reflected to the primary.

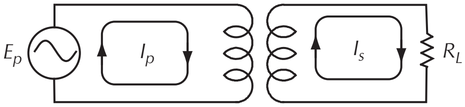


Figure 11-6. Cancellation of flux generated by load current.

Fig. 11-7 illustrates the relationships between voltage, excitation current, and flux in a transformer as *frequency* is changed. The horizontal scale is time. The primary voltage E_p is held constant as the frequency is changed (tripled and then tripled again). For example, the left waveform could represent one cycle at 100 Hz, the middle 300 Hz, and the right 900 Hz. Because of the primary inductance, excitation current I_p will decrease linearly with frequency—i.e., halving for every doubling in frequency or decreasing at 6 dB per octave. The magnitude of the magnetic flux will likewise *decrease* exactly the same way. Note that the inductance causes a 90° phase lag between voltage and current as well. Since the slew rate of a constant amplitude sine wave *increases* linearly with frequency—i.e., doubling for every doubling in frequency or increasing at 6 dB per octave—the resultant flux rate of change remains *constant*. Note that the slope of the I_p and flux waveforms stays constant as frequency is changed. Since, according to the law of induction, the voltage induced in the secondary is proportional to this slope or rate of change, output voltage also remains uniform, or flat versus frequency.

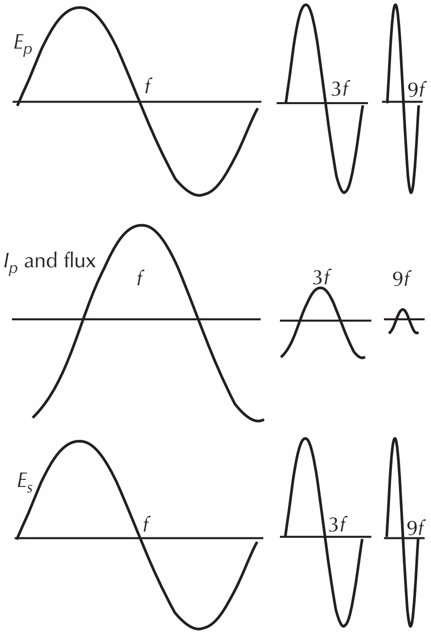


Figure 11-7. Excitation current and flux vary inversely with frequency.

11.1.2 Realities of Practical Transformers

Thus far, we have not considered the unavoidable parasitic elements which exist in any practical transformer. Even the design of a relatively simple 60 Hz power transformer must take parasitics into account. The design of an audio transformer operating over a 20 Hz to 20 kHz frequency range is much more difficult because these parasitics often interact in complex ways. For example, materials and techniques that improve low-frequency performance are often detrimental to high-frequency performance and vice versa. Good transformer designs must consider both the surrounding electronic circuitry and the performance ramifications of internal design tradeoffs.

A schematic representation of the major low frequency parasitic elements in a generalized transformer is shown in Fig. 11-8. The IDEAL TRANSFORMER represents a perfect transformer having a turns ratio of 1:*N* and no parasitic elements of any kind.

The actual transformer is connected at the PRI terminals to the driving voltage source, through its source impedance R_G, and at the SEC terminals to the load R_L.

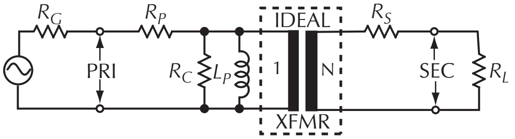


Figure 11-8. Transformer low-frequency parasitic elements.

One of the main goals in the design of any transformer is to reduce the excitation current in the primary winding to negligible levels so as not to become a significant load on the driving source. For a given source voltage and frequency, primary excitation current can be reduced only by increasing inductance L_P. In the context of normal electronic circuit impedances, very large values of inductance are required for satisfactory operation at the lowest audio frequencies. Of course, inductance can be raised by using a very large number of coil turns but, for reasons discussed later, there are practical limits due to other considerations. Another way to increase inductance by a factor of 10,000 or more is to wind the coil around a highly magnetic material, generally referred to as the *core.*

11.1.2.1 Core Materials and Construction

Magnetic circuits are quite similar to electric circuits. As shown in Fig. 11-11, magnetic flux always takes a closed path from one magnetic pole to the other and, like an electric current, always favors the paths of highest conductivity or least resistance. The equivalent of applied voltage in magnetic circuits is *magnetizing force*, symbolized H. It is directly proportional to ampere-turns (coil current I times its number of turns N) and inversely proportional to the flux path length ℓ in the magnetic circuit. The equivalent of electric current flow is *flux density*, symbolized B. It represents the number of magnetic flux lines per square unit of area. A graphic plot of the relationship between field intensity and flux density is shown in Fig. 11-9 and is referred to as the “B-H loop” or “hysteresis loop” for a given material. In the United States, the most commonly used units for magnetizing force and flux density are the Oersted and Gauss, respectively, which are CGS (centimeter, gram, second) units. In Europe, the SI (Systène Internationale) units amperes per meter and tesla, respectively, are more common. The slope of the B-H loop indicates how an incremental increase in applied magnetizing force changes the resulting flux density. This slope is effectively a measure of conductivity in the magnetic circuit and is called *permeability*, symbolized µ. Any material inside a coil, which can also serve as a form to support it, is called a *core*. By definition, the permeability of a vacuum, or air, is 1.00 and common nonmagnetic materials such as aluminum, brass, copper, paper, glass, and plastic also have a permeability of 1 for practical purposes. The permeability of some common ferromagnetic materials is about 300 for ordinary steel, about 5000 for 4% silicon transformer steel, and up to about 100,000 for some nickel-iron-molybdenum alloys. Because such materials concentrate magnetic flux, they greatly increase the inductance of a coil. Audio transformers must utilize both high-permeability cores and the largest practical number of coil turns to create high primary inductance. Coil inductance increases as the square of the number of turns and in direct proportion to the permeability of the core and can be approximated using the equation

$$L = \frac{3.2N^2\mu A}{10^8 l} \tag{11-4}$$

where,

L is the inductance in henrys,
N is the number of coil turns,
µ is the permeability of core,
A is the cross-section area of core in square inches,
l is the mean flux path length in inches.

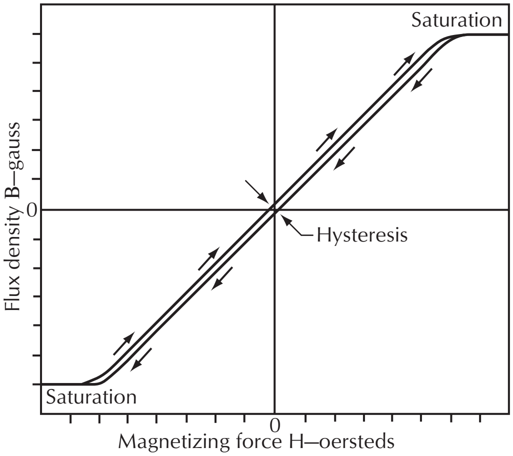


Figure 11-9. B-H loop for magnetic core material.

The permeability of magnetic materials varies with flux density. As shown in Fig. 11-9, when magnetic field intensity becomes high, the material can *saturate*, essentially losing its ability to conduct any additional flux. As a material saturates, its permeability decreases until, at complete saturation, its permeability becomes that of air or 1. In audio transformer applications, magnetic saturation causes low-frequency harmonic distortion to increase steadily for low-frequency signals as they increase in level beyond a threshold. In general, materials with a higher permeability tend to saturate at a lower flux density. In general, permeability also varies inversely with frequency.

Magnetic *hysteresis* can be thought of as a magnetic memory effect. When a magnetizing force saturates material that has high-hysteresis, it remains strongly magnetized even after the force is removed. High-hysteresis materials have wide or square B-H loops and are used to make magnetic memory devices and permanent magnets. However, if we magnetically saturate zero-hysteresis material, it will have no residual magnetism (flux density) when the magnetizing force is removed. But virtually all high-permeability core materials have *some* hysteresis, retaining a small memory of their previous magnetic state. Hysteresis can be greatly reduced by using certain metal alloys that have been annealed or heat-treated using special processes. In audio transformers, the nonlinearity due to magnetic hysteresis causes increased harmonic distortion for low-frequency signals at relatively low signal levels. Resistor R_C in Fig. 11-8 is a nonlinear resistance that, in the equivalent circuit model, represents the combined effects of magnetic saturation, magnetic hysteresis, and eddy-current losses.

The magnetic *operating point*, or zero signal point, for most transformers is the center of the B-H loop shown in Fig. 11-9, where the net magnetizing force is zero. Small ac signals cause a small portion of the loop to be traversed in the direction of the arrows. Large ac signals traverse portions farther from the operating point and may approach the saturation end points. For this normal operating point at the center, signal distortions (discussed in detail later) caused by the curvature of the loop are symmetrical—i.e., they affect the positive and negative signal excursions equally. Symmetrical distortions produce odd-order harmonics such as third and fifth. If dc current flows in a winding, the operating point will shift to a point on the loop away from the center. This causes the distortion of a superimposed ac signal to become nonsymmetrical. Nonsymmetrical distortions produce even-order harmonics such as second and fourth. When a small dc current flows in a winding, under say 1% of the saturation value, the effect is to add even-order harmonics to the normal odd-order content of the hysteresis distortion, which affects mostly low level signals. The same effects occur when the core becomes weakly magnetized, as could happen via the brief accidental application of dc to a winding, for example. However, the narrow B-H loop indicates that only a weak residual field would remain even if a magnetizing force strong enough to saturate the core were applied and then removed.

When a larger dc current flows in a winding, the symmetry of saturation distortion is also affected in a similar way. For example, enough dc current might flow in a winding to move the operating point to 50% of the core saturation value. Only half as much ac signal could then be handled before the core would saturate and, when it did, it would occur only for one direction of the signal swing. This would produce strong second-harmonic distortion. To avoid such saturation effects, *air gaps* are sometimes intentionally built into the magnetic circuit. This can be done, for example, by placing a thin paper spacer between the center leg of the E and I cores of Fig. 11-10. The magnetic permeability of such a gap is so low—even though it may be only a few thousandths of an inch—compared to the core material, that it effectively controls the flux density in the entire magnetic circuit. Although it drastically reduces the inductance of the coil, gapping is done to prevent flux density from reaching levels that would otherwise saturate the core, especially when substantial dc is present in a winding.

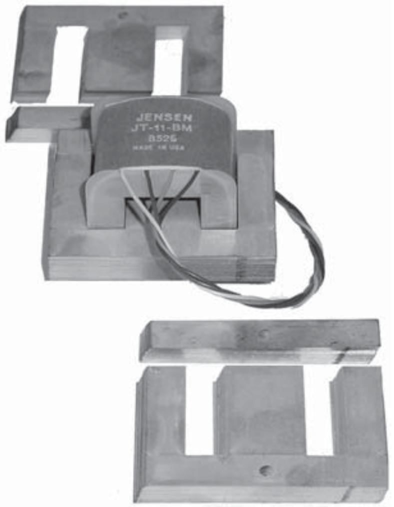


Figure 11-10. Core laminations are stacked and interleaved around bobbin that holds windings.

Because high-permeability materials are usually electrical conductors as well, small voltages are also induced in the cross-section of the core material itself, giving rise to *eddy currents*. Eddy currents are greatly reduced when the core consists of a stack of thin sheets called *laminations*, as shown in Fig. 11-10. Because the laminations are effectively insulated from each other, eddy currents generally become insignificant. The E and I shaped laminations shown form the widely used shell or double-window, core construction. Its parallel magnetic paths are illustrated in Fig. 11-11. When cores are made of laminations, care must be taken that they are flat and straight to avoid tiny air gaps between them that could significantly reduce inductance.

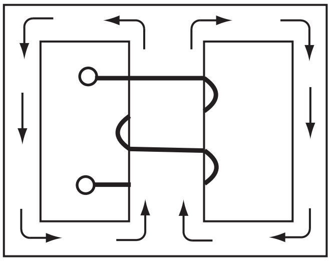


Figure 11-11. Magnetic circuits in shell core.

A *toroidal core* is made by rolling a long thin strip of core material into a coiled ring shape that looks something like a donut. It is insulated with a conformal coating or tape and windings are wound around the core through the center hole using special machines. With a toroidal core, there are no unintended air gaps that can degrade magnetic properties. Audio transformers don't often use toroidal cores because, especially in high bandwidth designs where multiple sections or Faraday shields are necessary, physical construction becomes very complex. Other core configurations include the *ring core*, sometimes called *semitoroidal*. It is similar to core of Fig. 11-11 but without the center section and windings are placed on the sides. Sometimes a solid—not laminations—metal version of a ring core is cut into two pieces having polished mating faces. These two *C-cores* are then held together with clamps after the windings are installed.

11.1.2.2 Winding Resistances and Auto-Transformers

If zero-resistance wire existed, some truly amazing transformers could be built. In a 60 Hz power transformer, for example, we could wind a primary with tiny wire on a tiny core to create enough inductance to make excitation current reasonable. Then we could wind a secondary with equally tiny wire. Because the wire has no resistance and the flux density in the core doesn't change with load current, this postage-stamp-sized transformer could handle unlimited kilowatts of power—and it wouldn't even get warm! But, at least until practical superconducting wire is available, real wire has resistance. As primary and secondary currents flow in the winding resistances, the resulting voltage drops cause signal loss in audio transformers and significant heating in power transformers. This resistance can be reduced by using larger—lower gauge—wire or fewer turns, but the required number of turns and the tolerable power loss (or resulting heat) all conspire to force transformers to become physically larger and heavier as their rated power increases. Sometimes silver wire is suggested to replace copper, but since its resistance is only about 6% less, its effect is minimal and certainly not cost-effective. However, there is an alternative configuration of transformer windings, called an *auto-transformer*, which can reduce the size and cost in certain applications. Because an auto-transformer electrically connects primary and secondary windings, it can't be used where electrical isolation is required! In addition, the size and cost advantage is maximum when the required turns ratio is very close to 1:1 and diminishes at higher ratios, becoming minimal in practical designs at about 3:1 or 1:3.

For example, in a hypothetical transformer to convert 100 V to 140 V, the primary could have 100 turns and the secondary 140 turns of wire. This transformer, with its 1:1.4 turns ratio, is represented in the upper diagram of Fig. 11-12. If 1 A of secondary (load) current I_S flows, transformer output power is 140 W and 1.4 A of primary current I_P will flow since input and output power must be equal in the ideal case. In a practical transformer, the wire size for each winding would be chosen to limit voltage losses and/or heating.

An auto-transformer essentially puts the windings in series so that the secondary voltage adds to (boosting) or subtracts from (bucking) the primary input voltage. A step-up auto-transformer is shown in the middle diagram of Fig. 11-12. Note that the dots indicate ends of the windings with the same instantaneous polarity. A 40 V secondary (the upper winding) series connected, as shown with the 100 V primary, would result in an output of 140 V. Now, if 1 A of secondary load current I_S flows, transformer output power is only 40 W and only 0.4 A of primary current I_P will flow. Although the total power delivered to the load is still 140 W, 100 W have come directly from the driving source and only 40 W have been transformed and added by the auto-transformer. In the auto-transformer, 100 turns of smaller wire can be used for the primary and only 40

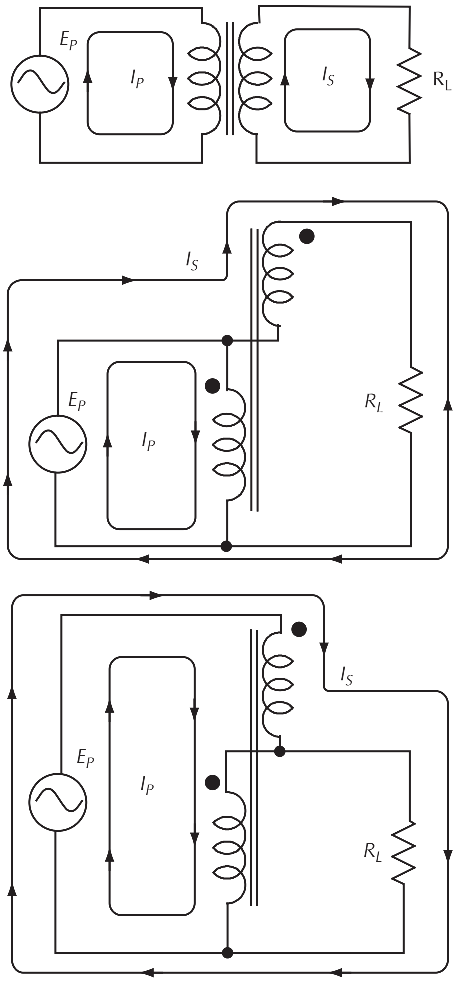


Figure 11-12. Auto-transformers employ a buck/boost principle.

turns of heavier wire is needed for the secondary. Compare this to the total of 240 turns of heavier wire required in the transformer.

A step-down auto-transformer is shown in the bottom diagram of Fig. 11-12. Operation is similar except that the secondary is connected so that its instantaneous polarity subtracts from or bucks the input voltage. For example, we could step down U.S. 120 Vac power to Japanese 100 Vac power by configuring a 100 V to 20 V step-down transformer as an auto-transformer. Thus, a 100 W load can be driven using only a 20 W rated transformer.

The windings of low level audio transformers may consist of hundreds or even many thousands of turns of wire, sometimes as small as #46 gauge, whose 0.0015 inch diameter is comparable to a human hair. As a result, each winding may have a dc resistance as high as several thousand ohms. Transformer primary and secondary winding resistances are represented by R_P and R_S, respectively, in Fig. 11-8.

11.1.2.3 Leakage Inductance and Winding Techniques

In an ideal transformer, since all flux generated by the primary is linked to the secondary, a short circuit on the secondary would be reflected to the primary as a short circuit. However, in real transformers, the unlinked flux causes a residual or *leakage inductance* that can be measured at either winding. Therefore, the secondary would appear to have residual inductance if the primary were shorted and vice-versa. The leakage inductance is shown as L_L in the model of Fig. 11-13. Note that leakage inductance is reflected from one winding to another as the square of turns ratio, just as other impedances are.

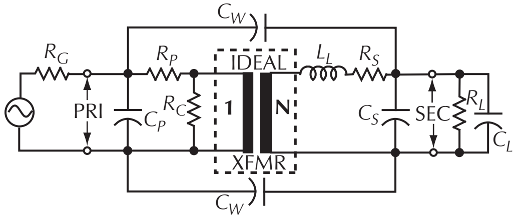


Figure 11-13. Transformer high frequency parasitic elements.

The degree of flux coupling between primary and secondary windings depends on the physical spacing between them and how they are placed with respect to each other. The lowest leakage inductance is achieved by winding the coils on a common axis and as close as possible to each other. The ultimate form of this technique is called *multi-filar* winding where multiple wires are wound simultaneously as if they were a single strand. For example, if two windings—i.e. primary and secondary—are wound as one, the transformer is said to be *bi-filar* wound. Note in the cross-section view of Fig. 11-14 how the primary and secondary windings are side-by-side throughout the entire winding. Another technique to reduce leakage inductance is to use *layering*, a technique in which portions or *sections* of the primary and/or secondary are wound in sequence over each other to interleave them. For example, Fig. 11-15 shows the cross-section of a three-layer transformer where half the primary is wound, then the secondary, followed by the other half of the primary. This results in considerably less leakage inductance than just a secondary over primary two-layer design. Leakage inductance decreases rapidly as the number of layers is increased.

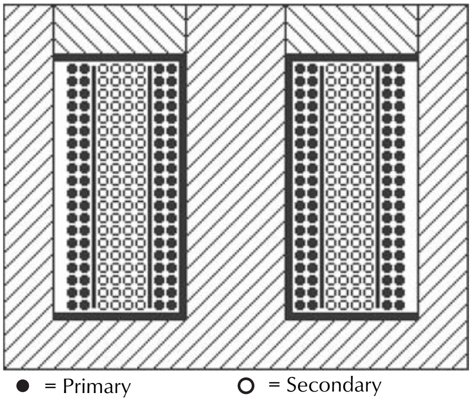


Figure 11-14. Layered windings.

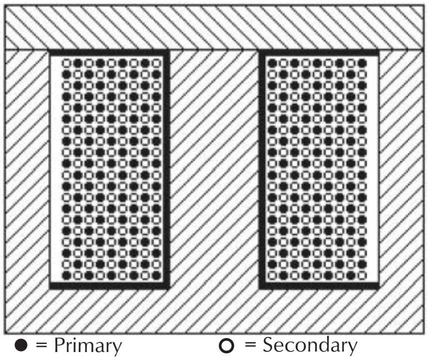


Figure 11-15. Bi-filar windings.

11.1.2.4 Winding Capacitances and Faraday Shields

To allow the maximum number of turns in a given space, the insulation on the wire used to wind transformers is very thin. Called *magnet wire*, it is most commonly insulated by a thin film of polyurethane enamel. A transformer winding is made, in general, by spinning the bobbin shown in Fig. 11-10 on a machine similar to a lathe and guiding the wire to form a layer one wire thick across the length of the bobbin. The wire is guided to traverse back and forth across the bobbin to form a coil of many layers as shown in Fig. 11-15, where the bobbin cross-section is the solid line on three sides of the winding. This simple side-to-side, back-and-forth winding results in considerable layer-to-layer capacitance within a winding or winding section. More complex techniques such as universal winding are sometimes used to substantially reduce winding capacitances. These capacitances within the windings are represented by C_P and C_S in the circuit model of Fig. 11-13. Additional capacitances will exist between the primary and secondary windings and are represented by capacitors C_W in the model. Sometimes layers of insulating tape are added to increase the spacing, therefore reducing capacitance, between primary and secondary windings. In the bi-filar windings of Fig. 11-14, since the wires of primary and secondary windings are side by side throughout, the inter-winding capacitances C_W can be quite high.

In some applications, interwinding capacitances are very undesirable. They are completely eliminated by the use of a *Faraday shield* between the windings. Sometimes called an *electrostatic shield*, it generally takes the form of a thin sheet of copper foil placed between the windings. Obviously, transformers that utilize multiple layers to reduce leakage inductance will require Faraday shields between all adjacent layers. In Fig. 11-15 the dark lines between the winding layers are the Faraday shields. Normally, all the shields surrounding a winding are tied together and treated as a single electrical connection. When connected to circuit ground, as shown in Fig. 11-16, a Faraday shield intercepts the capacitive current that would otherwise flow between transformer windings.

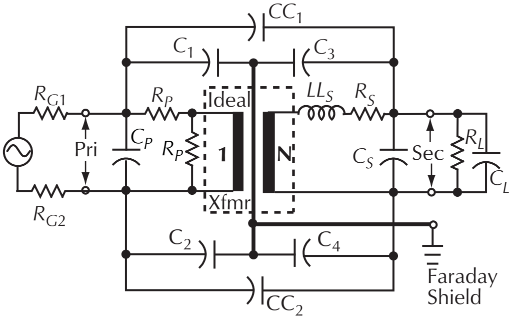


Figure 11-16. High frequency equivalent circuit of a transformer with Faraday shield and driven by a balanced source.

Faraday shields are nearly always used in transformers designed to eliminate ground noise. In these applications, the transformer is intended to respond only to the voltage difference or signal across its primary and have no response to the noise that exists equally (or common-mode) at the terminals of its primary. A Faraday shield is used to prevent capacitive coupling, via C_W in Fig. 11-13, of this noise to the secondary. For any winding connected to a balanced line, the matching of capacitances to ground is critical to the rejection of common-mode noise, or CMRR, as discussed in Chapter 37. In Fig. 11-16, if the primary is driven by a balanced line, C_1 and C_2 must be very accurately matched to achieve high CMRR. In most applications, such as microphone or line input transformers, the secondary is operated unbalanced—i.e., one side is grounded. This relaxes the matching requirements for capacitances C_3 and C_4. Although capacitances CC_1 and CC_2 are generally quite small—a few pF—they have the effect of diminishing CMRR at high audio frequencies and limiting rejection of RF interference.

11.1.2.5 Magnetic Shielding

A *magnetic shield* has a completely different purpose. Devices such as power transformers, electric motors, and television or computer monitor cathode-ray tubes generate powerful ac magnetic fields. If such a field takes a path through the core of an audio transformer, it can induce an undesired voltage in its windings—most often heard as hum. If the offending source and the victim transformer have fixed locations, orientation of one or both can sometimes nullify the pickup. In Fig. 11-11 note that an external field that flows vertically through the core will cause a flux gradient across the length of the coil, inducing a voltage in it, but a field that flows horizontally through the core will not. Such magnetic pickup is usually worse in input transformers (discussed later) because they generally have more turns. It should also be noted that higher permeability core materials are more immune to external fields. Therefore, an unshielded output transformer with a high nickel core will be more immune than one with a steel core.

Another way to prevent such pickup is to surround the core with a closed—no air gap—magnetic path. This *magnetic shield* most often takes the form of a can or box with tight-fitting lid and is made of high permeability material. While the permeability of ordinary steel, such as that in electrical conduit, is only about 300, special-purpose nickel alloys can have permeability as high as 100,000. Commercial products include Mumetal®, Permalloy®, HyMu® and Co-Netic®.[1,2] Since the shield completely surrounds the transformer, the offending external field will now flow through it instead of the transformer core. Generally speaking, care must be taken not to mechanically stress these metals because doing so will significantly decrease their permeability. For this reason, most magnetic shield materials must be re-annealed after they are fabricated.

The effectiveness of magnetic shielding is generally rated in dB. The transformer is placed in an external magnetic field of known strength, generally at 60 Hz. Its output without and with the shield is then compared. For example, a housing of $^1/_8$ inch thick cast iron reduces pickup by about 12 dB, while a 0.030 inch thick Mumetal can reduces it by about 30 dB. Where low-level transformers operate near strong magnetic fields, several progressively smaller shield cans can be nested around the transformer. Two or three Mumetal cans can provide 60 dB and 90 dB of shielding, respectively. In very strong fields, because high permeability materials might saturate, an iron or steel outer can is sometimes used.

Toroidal power transformers can have a weaker radiated magnetic field than other types. Using them can be an advantage if audio transformers must be located nearby. However, a toroidal transformer must be otherwise well designed to produce a low external field. For example, every winding must completely cover the full periphery of the core. The attachment points of the transformer lead wires are frequently a problem in this regard. To gain size and cost advantages, most commercial power transformers of any kind are designed to operate on the verge of magnetic saturation of the core. When saturation occurs in any transformer, magnetic field radiation dramatically increases. Power transformers designed to operate at low flux density will prevent this. A standard commercial power transformer, when operated at reduced primary voltage, will have a very low external field—comparable to that of a standard toroidal design.

11.1.3 General Application Considerations

For any given application, a number of parameters must be considered when selecting or designing an appropriate audio transformer. We will discuss how the performance of a transformer can be profoundly affected by its interaction with surrounding circuitry.

11.1.3.1 Maximum Signal Level, Distortion, and Source Impedance

Because these parameters are inextricably interdependent, they must be discussed as a group. Although transformer operating level is often specified in terms of power such as dBm or watts, what directly affects distortion is the equivalent driving voltage. Distortion is caused by excitation current in the primary winding which is proportional to primary voltage, not power. Referring to Fig. 11-8, recall that nonlinear resistance R_C represents the distortion-producing mechanisms of the core material. Consider that, if both R_G, the driving source impedance, and R_P, the internal winding resistance, were zero, the voltage source—by definition zero impedance—would effectively short out R_C, resulting in zero distortion! But in a real transformer design there is a fixed relationship between signal level, distortion, and source impedance. Since distortion is also a function of magnetic flux density, which increases as frequency *decreases*, a maximum operating level specification must also specify a frequency. The specified maximum operating level, maximum allowable distortion at a specified low frequency, and maximum allowable source impedance will usually dictate the type of core material that must be used and its physical size. And, of course, cost plays a role, too.

The most commonly used audio transformer core materials are *M6 steel* (a steel alloy containing 6% silicon) and 49% *nickel* or 84% *nickel* (alloys containing 49% or 84% nickel plus iron and molybdenum). Nickel alloys are substantially more expensive than steel. Fig. 11-17 shows how the choice of core material affects low-frequency distortion as signal level changes. The increased distortion at low levels is due to magnetic hysteresis and at high levels is due to magnetic saturation. Fig. 11-18 shows how distortion decreases rapidly with increasing frequency. Because of differences in their hysteresis distortion, the falloff is most rapid for the 84% nickel and least rapid for the steel. Fig. 11-19 shows how distortion is strongly affected by the impedance of the driving source. The plots begin at 40 Ω because that is the resistance of the primary winding. Therefore, maximum operating levels predicated on higher frequencies, higher distortion, and lower source impedance will always be higher than those predicated on lower frequencies, lower distortion, and lower source impedance.

As background, it should be said that THD, or total harmonic distortion, is a remarkably inadequate way to describe the perceived awfulness of distortion. Distortion consisting of low-order harmonics, 2nd or 3rd for example, is dramatically less audible than that consisting of high order harmonics, 7th or 13th for example. Consider that, at very low frequencies, even the finest loudspeakers routinely exhibit harmonic distortion in the range of several percent at normal listening levels. Simple distortion tests whose results correlate well with the human auditory experience simply don't exist. Clearly, such perceptions are far too complex to quantify with a single figure.

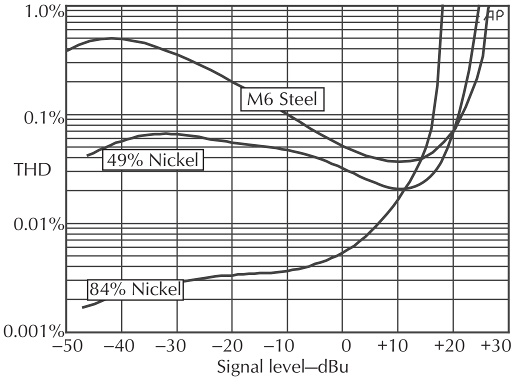


Figure 11-17. Measured THD at 20 Hz and 40 Ω source versus signal level for three types of core materials.

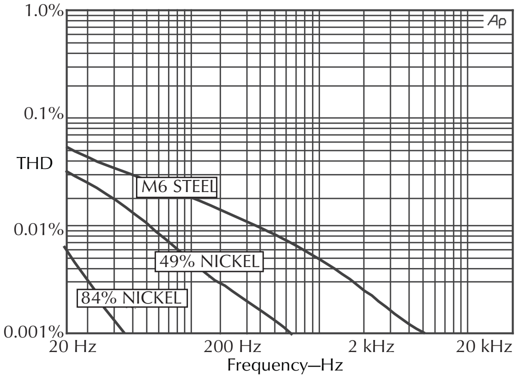


Figure 11-18. Measured THD at 0 dBu and 40 Ω source versus frequency for the cores of Figure 11-17.

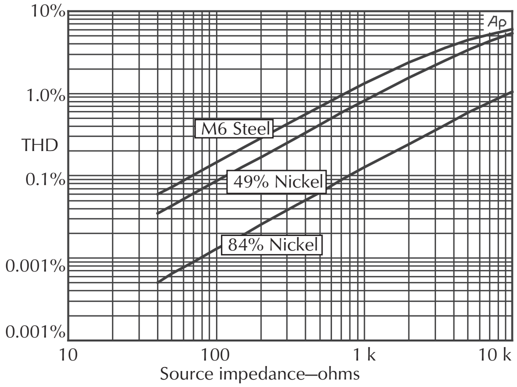


Figure 11-19. Measured THD at 0 dBu and 20 Hz versus source impedance for the cores of Figs. 11-17 and 11-18.

One type of distortion that is particularly audible is intermodulation or IM distortion. Test signals generally combine a large low-frequency signal with a smaller high-frequency signal and measure how much the amplitude of the high frequency is modulated by the lower frequency. Such intermodulation creates tones at new, nonharmonic frequencies. The classic SMPTE (Society of Motion Picture and Television Engineers)

IM distortion test mixes 60 Hz and 7 kHz signals in a 4:1 amplitude ratio. For virtually all electronic amplifier circuits, there is an approximate relationship between harmonic distortion and SMPTE IM distortion. For example, if an amplifier measured 0.1% THD at 60 Hz at a given operating level, its SMPTE IM distortion would measure about three or four times that, or 0.3% to 0.4% at an equivalent operating level. This correlation is due to the fact that electronic non-linearities generally distort audio signals without regard to frequency. Actually, because of negative feedback and limited gain bandwidth, most electronic distortions become worse as frequency increases.

Distortion in audio transformers is different in a way that makes it sound unusually benign. It is caused by the smooth symmetrical curvature of the magnetic transfer characteristic or B-H loop of the core material shown in Fig. 11-9. The nonlinearity is related to flux density that, for a constant voltage input, is inversely proportional to frequency. The resulting harmonic distortion products are nearly pure third harmonic. In Fig. 11-18, note that distortion for 84% nickel cores roughly quarters for every doubling of frequency, dropping to less than 0.001% above about 50 Hz. Unlike that in amplifiers, the distortion mechanism in a transformer is frequency selective. This makes its IM distortion much less than might be expected. For example, the Jensen JT-10KB-D line input transformer has a THD of about 0.03% for a +26 dBu input at 60 Hz. But, at an equivalent level, its SMPTE IM distortion is only about 0.01%—about a tenth of what it would be for an amplifier having the same THD.

11.1.3.2 Frequency Response

The simplified equivalent circuit of Fig. 11-20 shows the high-pass RL filter formed by the circuit resistances and transformer primary inductance L_P. The effective source impedance is the parallel equivalent of $R_G + R_P$ and $R_S + R_L$. When the inductive reactance of L_P equals the effective source impedance, low-frequency response will fall to 3 dB below its mid-band value. For example, consider a transformer having an L_P of 10 henrys and winding resistances R_P and R_S of 50 Ω each. The generator impedance R_G is 600 Ω and the load R_L is 10 kΩ. The effective source impedance is then 600 Ω + 50 Ω in parallel with 10 kΩ + 50 Ω, which computes to about 610 Ω. A 10 henry inductor will have 610 Ω of reactance at about 10 Hz, making response 3 dB down at that frequency. If the generator impedance R_G were made 50 Ω instead, response would be −3 dB at 1.6 Hz. Lower source impedance will always extend low-frequency bandwidth. Since the filter is single pole, response falls at 6 dB per octave. As discussed earlier, the permeability of most core material steadily increases as frequency is lowered and typically reaches its maximum somewhere under 1 Hz. This results in an actual roll-off rate *less* than 6 dB per octave and a corresponding improvement in phase distortion—deviation from linear phase. Although a transformer cannot have response to 0 Hz or dc, it can have much less phase distortion than a coupling capacitor chosen for the same cutoff frequency. Or, as a salesperson might say, "It's not a defect, it's a feature."

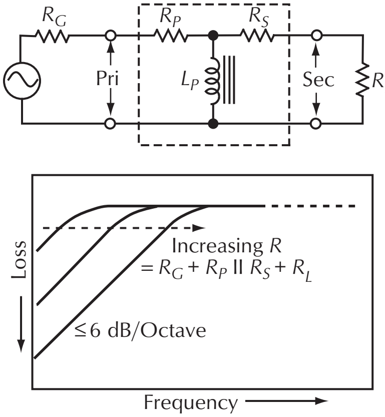


Figure 11-20. Simplified low frequency transformer equivalent circuit.

The simplified equivalent schematic of Fig. 11-21 shows the parasitic elements that limit and control high-frequency response.

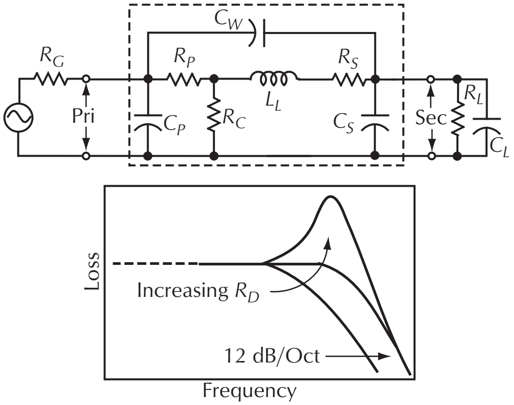


Figure 11-21. Simplified high-frequency transformer equivalent circuit.

Except in bi-filar wound types discussed below, leakage inductance L_L and load capacitance are the major limiting factors. This is especially true in Faraday shields because of the increase in leakage inductance. Note that a low-pass filter is formed by series leakage inductance L_L with shunt winding capacitance C_S plus external load capacitance C_L. Since this filter has two reactive elements, it is a two-pole filter subject to response variations caused by *damping*. Resistive elements in a filter provide damping, dissipating energy when the inductive and capacitive elements resonate. As shown in the figure, if damping resistance R_D is too high, response will rise before it falls and if damping resistance is too low, response falls too early. Optimum damping results in the widest bandwidth with no response peak. It should be noted that placing capacitive loads C_L on transformers with high leakage inductance not only lowers their bandwidth but also changes the resistance required for optimum damping. For most transformers, R_L controls damping. In the time domain, under-damping manifests itself as ringing on square-waves as shown in Fig. 11-22. When loaded by its specified load resistance, the same transformer responds as shown in Fig. 11-23. In some transformers, source impedance also provides significant damping.

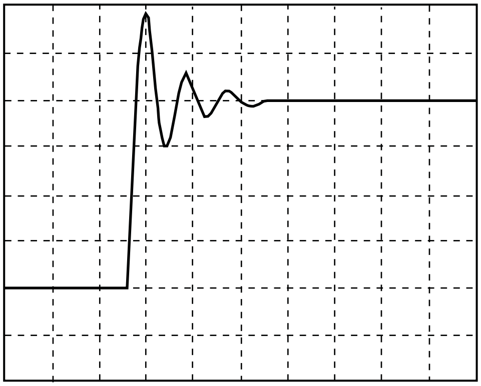


Figure 11-22. Undamped response.

In bi-filar wound transformers, leakage inductance L_L is very low but interwinding capacitance C_W and winding capacitances C_P and C_S are quite high. Leakage inductance must be kept very small in applications such as line drivers because large cable capacitances C_L would otherwise be disastrous to high-frequency response. Such transformers are generally referred to as output transformers. Also note that a low-pass filter is formed by series R_G and shunt C_P plus C_S. Therefore, driving sources may limit high-frequency response if their source impedance R_G is too high. In normal 1:1 bi-filar output transformer designs, C_W actually works to capacitively couple very high frequencies between windings. Depending on the application, this can be either a defect or a feature.

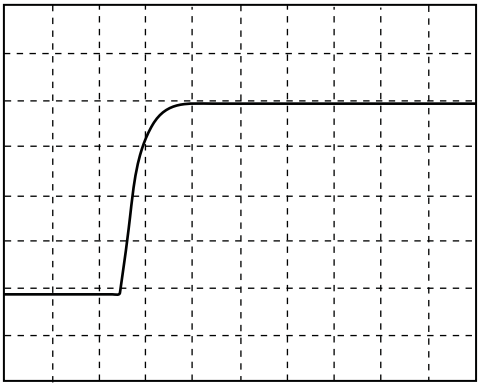


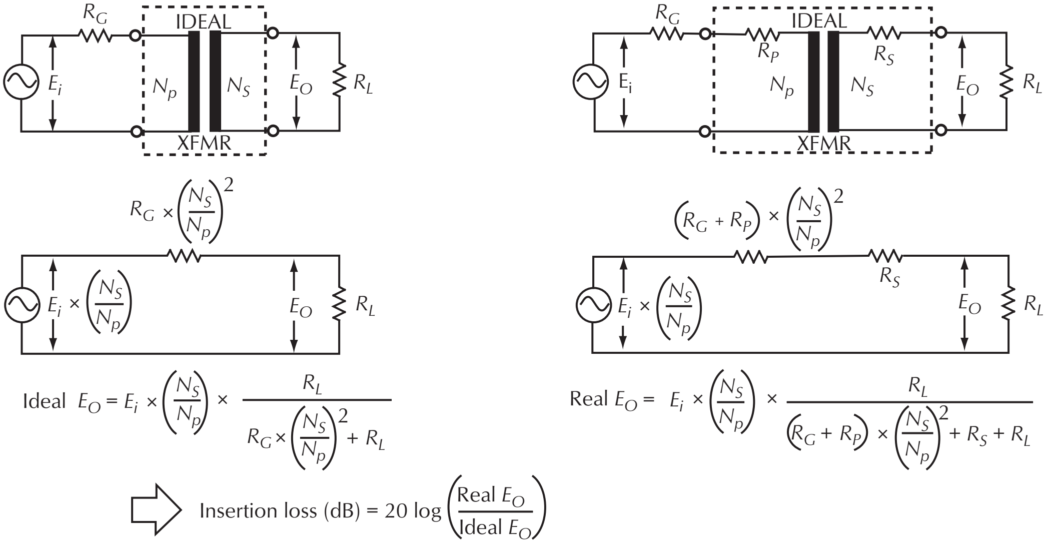
Figure 11-23. Proper damping.

11.1.3.3 Insertion Loss

The power output from a transformer will always be slightly less than power input to it. As current flows in its windings, their dc resistance causes additional voltage drops and power loss as heat. Broadly defined, *insertion loss* or *gain* is that caused by inserting a device into the signal path. But, because even an ideal lossless transformer can increase or decrease signal level by virtue of its turns ratio, the term *insertion loss* is usually defined as the difference in output signal level between the real transformer and an ideal one with the same turns ratio.

The circuit models, Thevenin equivalent circuits, and equations for both ideal and real transformers are shown in Fig. 11-24. For example, consider an ideal 1:1 turns ratio transformer and $R_G = R_L = 600\ \Omega$. Since N_s/N_p is 1, the equivalent circuit becomes simply E_i in series with R_G or 600 Ω. When R_L is connected, a simple voltage divider is formed, making $E_O = 0.5\ E_i$ or a 6.02 dB loss. For a real transformer having $R_P = R_S = 50\ \Omega$, the equivalent circuit becomes E_i in series with $R_G + R_P + R_S$ or 700 Ω. Now, the output $E_O = 0.462\ E_i$ or a 6.72 dB loss. Therefore, the insertion loss of the transformer is 0.70 dB.

Calculations are similar for transformers with turns ratios other than 1:1, except that voltage is multiplied by the turns ratio and reflected impedances are multiplied by the turns ratio squared as shown in the equations. For example, consider a 2:1 turns ratio transformer, $R_G = 600\ \Omega$, and $R_L = 150\ \Omega$. The ideal transformer


R_G
IDEAL
E_i
N_P
N_S
E_O
R_L
XFMR
R_P
R_S
$R_G \times \left(\frac{N_S}{N_P}\right)^2$
$(R_G + R_P) \times \left(\frac{N_S}{N_P}\right)^2$
$E_i \times \left(\frac{N_S}{N_P}\right)$

$$\text{Ideal } E_O = E_i \times \left(\frac{N_S}{N_P}\right) \times \frac{R_L}{R_G \times \left(\frac{N_S}{N_P}\right)^2 + R_L}$$

$$\text{Real } E_O = E_i \times \left(\frac{N_S}{N_P}\right) \times \frac{R_L}{(R_G + R_P) \times \left(\frac{N_S}{N_P}\right)^2 + R_S + R_L}$$

$$\text{Insertion loss (dB)} = 20 \log \left(\frac{\text{Real } E_O}{\text{Ideal } E_O}\right)$$

Figure 11-24. Insertion loss compares the outputs of real and ideal transformers.

output appears as 0.5 E_i in series with $R_G/4$ or 150 Ω. When R_L is connected, a simple voltage divider is formed making $E_O = 0.25\ E_i$ or a 12.04 dB loss. For a real transformer having $R_P = 50\ \Omega$ and $R_S = 25\ \Omega$, the equivalent circuit becomes 0.5 E_i in series with $(R_G + R_P)/4 + R_S$ or 187.5 Ω. Now, the output $E_O = 0.222\ E_i$ or a 13.07 dB loss. Therefore, the insertion loss of this transformer is 1.03 dB.

11.1.3.4 Sources with Zero Impedance

One effect of using negative feedback around a high gain amplifier is to reduce its output impedance. Output impedance is reduced by the feedback factor, which is open-loop gain in dB minus closed-loop gain in dB. A typical op-amp with an open-loop gain of 80 dB, set for closed-loop gain of 20 dB, the feedback factor is 80 dB − 20 dB = 60 dB or 1000, will have its open-loop output impedance of 50 Ω reduced by the feedback factor (1000) to about 0.05 Ω. Within the limits of linear operation—i.e., no current limiting or voltage clipping—the feedback around the amplifier effectively forces the output to remain constant regardless of loading. For all practical purposes the op-amp output can be considered a true voltage source.

As seen in Fig. 11-19, the distortion performance of any transformer is significantly improved when the driving source impedance is less than the dc resistance of the primary. However, little is gained for source impedances below about 10% of the winding dc resistance. For example, consider a typical line output transformer with a primary dc resistance of 40 Ω. A driving source impedance well under 4 Ω will result in lowest distortion. The line drivers shown in Fig. 11-28 and Fig. 11-29 use a paralleled inductor and resistor to isolate or decouple the amplifier from the destabilizing effects of load (cable) capacitance at very high frequencies. Because the isolator's impedance is well under an ohm at all audio frequencies, it is much preferred to the relatively large series, or build-out, resistor often used for the purpose. It is even possible for an amplifier to generate *negative* output resistance to cancel the winding resistance of the output transformer. Audio Precision uses such a patented circuit in their System 1 audio generator to reduce transformer-related distortion to extremely low levels.

11.1.3.5 Bi-Directional Reflection of Impedances

The impedances associated with audio transformers may seem confusing. Much of the confusion probably stems from the fact that *transformers can simultaneously reflect two different impedances—one in each direction*. One is the impedance of the driving source, as seen from the secondary, and the other is the impedance of the load, as seen from the primary. Transformers simply reflect impedances, modified by the square of their turns ratio, from one winding to another. However, because of their internal parasitic elements discussed previously, transformers tend to produce optimum results when used within a specified range of external impedances.

There is essentially no intrinsic impedance associated with the transformer itself. With no load on its secondary, the primary of a transformer is just an inductor and its impedance will vary linearly with frequency. For example, a 5 H primary winding would have an input impedance of about 3 kΩ at 100 Hz, 30 kΩ at 1 kHz, and 300 kΩ at 10 kHz. In a proper transformer design, this self-impedance, as well as those of other internal parasitics, should have negligible effects on normal circuit operation. The following applications will illustrate the point.

A 1:1 *output* transformer application is shown in Fig. 11-25. It has a winding inductance of about 25 H and negligible leakage inductance. The open circuit impedance, at 1 kHz, of either winding is about 150 kΩ. Since the dc resistance is about 40 Ω per winding, if the primary is short circuited, the secondary impedance will drop to 80 Ω. If we place the transformer between a zero-impedance amplifier (more on that later) and a load, the amplifier will see the load through the transformer and the load will see the amplifier through the transformer. In our example, the amplifier would look like 80 Ω to the output line/load and the 600 Ω line/load would look like 680 Ω to the amplifier. If the load were 20 kΩ, it would look like slightly less than 20 kΩ because the open circuit transformer impedance, 150 kΩ at 1 kHz, is effectively in parallel with it. For most applications, these effects are trivial.

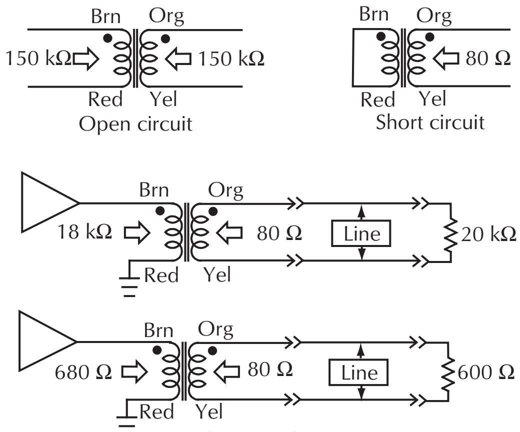


Figure 11-25. Impedance reflection in a 1:1 transformer.

A 4:1 *input* transformer example is shown in Fig. 11-26. It has a primary inductance of about 300 H and negligible winding capacitance. The open circuit impedance, at 1 kHz, of the primary is about 2 MΩ. Because this transformer has a 4:1 turns ratio and, therefore a 16:1 impedance ratio, the secondary open circuit impedance is about 125 kΩ. The dc resistances are about 2.5 kΩ for the primary and 92 Ω for the secondary. Since this is an input transformer, it must be used with the specified secondary load resistance of 2.43 kΩ for proper damping (flat frequency response). This load on the secondary will be transformed by the turns ratio to look like about 42 kΩ at the primary. To minimize the noise contribution of the amplifier stage, we need to know what the transformer secondary looks like, impedancewise, to the amplifier. If we assume that the primary is driven from the line in our previous output transformer example with its 80 Ω source impedance, we can calculate that the secondary will look like about 225 Ω to the amplifier input. Actually, any source impedance less than 1 kΩ would have little effect on the impedance seen at the secondary.

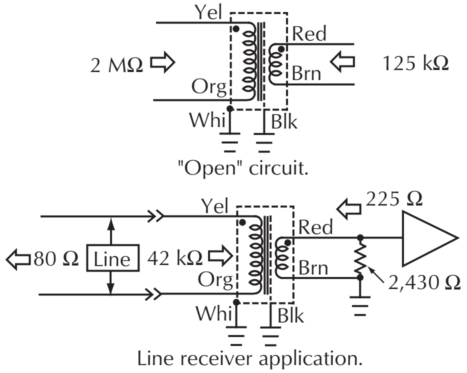


Figure 11-26. Impedance reflection in a 4:1 transformer.

Transformers are not intelligent—they can't magically couple signals in one direction only. Magnetic coupling is truly bi-directional. For example, Fig. 11-27 shows a three-winding 1:1:1 transformer connected to drive two 600 Ω loads. The driver sees the loads in parallel or, neglecting winding resistances, 300 Ω. Likewise, a short on either output will be reflected to the driver as a short. Of course, turns ratios and winding resistances must be taken into account to calculate actual driver loading. For the same reason, stereo L and R outputs that drive two windings on the same transformer are effectively driving each other, possibly causing distortion or damage.

11.1.3.6 Transformer Noise Figure

Although the step-up turns ratio of a transformer may provide noise-free voltage gain, some 20 dB for a 1:10 turns ratio, it's important to understand that improve-

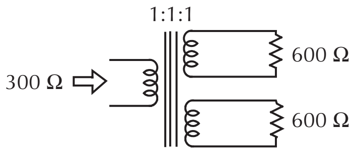
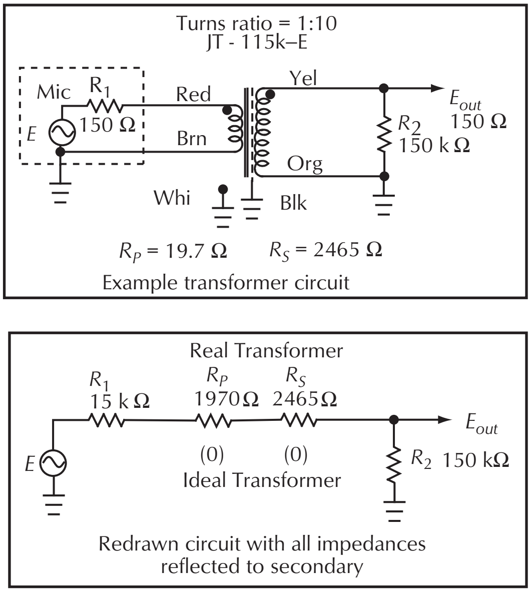


Figure 11-27. Multiple loads are effectively paralleled.

ments in signal-to-noise ratio are not solely due to this gain. Because most amplifying devices generate current noise as well as voltage noise at their inputs, their noise performance will suffer when turns ratio is not the optimum for that particular amplifier (see 21.1.2.3 Microphone Preamp Noise). *Noise figure* measures, in dB, how much the output signal-to-noise ratio of a system is degraded by a given system component. All resistances, including the winding resistances of transformers, generate thermal noise. Therefore, the noise figure of a transformer indicates the increase in thermal noise or hiss when it replaces an ideal noiseless transformer having the same turns ratio—i.e., voltage gain. The noise figure of a transformer is calculated as shown in Fig. 11-28.

11.1.3.7 Basic Classification by Application

Many aspects of transformer performance, such as level handling, distortion, and bandwidth, depend critically on the impedance of the driving source and, in most cases, the resistance and capacitance of the load. These impedances play such an important role that they essentially classify audio transformers into two basic types. Most simply stated, *output* transformers are used when load impedances are low, as in line drivers, while *input* transformers are used when load impedances are high, as in line receivers. The load for a line driver is not just the high-impedance equipment input it drives—it also includes the cable capacitance, whose impedance can become quite low at 20 kHz. The conflicting technical requirements for output and input types make their



Primary side impedances get multiplied by the square of the turns ratio.

$R_1 = 150\ \Omega \times 10^2 = 15\ \text{k}\Omega$
$R_P = 19.7\ \Omega \times 10^2\ \Omega = 1970\ \Omega$
R_S = Secondary DCR = 2465 Ω
$R_2 = 150\ \text{k}\Omega$ (load)

The transformer noise figure is calculated by comparing a real transformer with its winding resistances to an ideal transformer with no winding resistances.
First, transform all impedances to the secondary as shown to the left.

There are two components to the calculation.

1. The additional noise due to the increased output impedance.

$$\text{REAL } Z_{out} \quad \frac{150\text{ k} \times (15\text{ k} + 1970 + 2465)}{150\text{ k} + 15\text{ k} + 1970 + 2465} = 17.205\text{ k}\Omega$$

$$\text{IDEAL } Z_{out} \quad \frac{1150\text{ k} + 15\text{ k}}{50\text{ k} \times 15\text{ k}} = 13.636\text{ k}\Omega$$

$$NF = 20\text{Log}\sqrt{\frac{17{,}206\ (REAL)}{13{,}636\ (IDEAL)}} = 1.01\text{ dB}$$

2. The decrease in signal level at the output due to the increased series losses.

$$IDEAL\ E_{out} \quad \frac{150\text{ k}}{150\text{ k} + 15\text{ k}} = 0.909$$

$$REAL_{out} \quad \frac{150\text{ k}}{150\text{ k} + 15\text{ k} + 1970 + 2465} = 0.885$$

$$NF = 20\text{Log}\frac{0.909\ (IDEAL)}{0.885\ (REAL)} = 0.232\text{ dB}$$

3. *Total NF* = 1.01 dB + 0.232 dB = 1.23 dB

Figure 11-28. Finding the noise figure of a transformer.

design and physical construction very different. Of course, some audio transformer applications need features of both input and output transformers and are not so easily classified.

Output transformers must have very low leakage inductance in order to maintain high-frequency bandwidth with capacitive loads. Because of this, they rarely use Faraday shields and are most often multi-filar wound. For low insertion loss, they use relatively few turns of large wire to decrease winding resistances. Since they use fewer turns and operate at relatively high signal levels, output transformers seldom use magnetic shielding. On the other hand, input transformers directly drive the usually high-resistance, low-capacitance input of amplifier circuitry. Many input transformers operate at relatively low signal levels, frequently have a Faraday shield, and are usually enclosed in at least one magnetic shield.

11.2 Audio Transformers for Specific Applications

Broadly speaking, audio transformers are used because they have two very useful properties. First, they can benefit circuit performance by transforming circuit impedances, to optimize amplifier noise performance, for example. Second, because there is no direct electrical connection between its primary and secondary windings, a transformer provides electrical or galvanic *isolation* between two circuits. As discussed in Chapter 37, isolation in signal circuits is a powerful technique to prevent or cure noise problems caused by normal ground voltage differences in audio systems. To be truly useful, a transformer should take full advantage of one or both of these properties but not compromise audio performance in terms of bandwidth, distortion, or noise.

11.2.1 Equipment-Level Applications

11.2.1.1 Microphone Input

A microphone input transformer is driven by the nominal 150 Ω, or 200 Ω in Europe, source impedance of professional microphones. One of its most important functions is to transform this impedance to a generally higher one more suited to optimum noise performance. As discussed in Chapter 21, this optimum impedance may range from 500 Ω to over 15 kΩ, depending on the amplifier. For this reason, microphone input transformers are made with turns ratios ranging from 1:2 to 1:10 or higher. The circuit of Fig. 11-29 uses a 1:5 turns ratio transformer, causing the microphone to appear as a 3.7 kΩ source to the IC amplifier, which optimizes its noise. The input impedance of the transformer is about 1.5 kΩ. It is important that this impedance remain reasonably flat with frequency to avoid altering the microphone response at frequency extremes, see Fig. 21-6.

In all balanced signal connections, common-mode noise can exist due to ground voltage differences or magnetic or electrostatic fields acting on the interconnecting cable. It is called common mode noise because it appears equally on the two signal lines, at least in theory. Perhaps the most important function of a balanced input is to reject (not respond to) this common-mode noise. A figure comparing the ratio of its differential or normal signal response to its common-mode response is called *common mode rejection ratio* or CMRR. An input transformer must have two attributes to achieve high CMRR. First, the capacitances of its two inputs to ground must be very well matched and as low as possible. Second, it must have minimal capacitance between its primary and secondary windings. This is usually accomplished by precision winding of

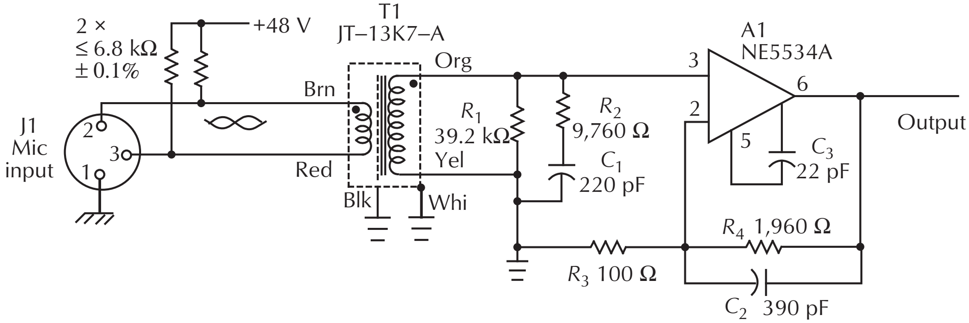


Figure 11-29. Microphone preamplifier with 40 dB overall gain.

the primary to evenly distribute capacitances and the incorporation of a Faraday shield between primary and secondary. Because the common-mode input impedances of a transformer consist only of capacitances of about 50 pF, transformer CMRR is maintained in real-world systems where the source impedances of devices driving the balanced line and the capacitances of the cable itself are not matched with great precision.[3]

Because tolerable common-mode voltage is limited only by winding insulation, transformers are well suited for phantom power applications. The standard arrangement using precision resistors is shown in Fig. 11-29. Resistors of lesser precision may degrade CMRR. Feeding phantom power through a center tap on the primary requires that both the number of turns and the dc resistance on either side of the tap be precisely matched to avoid small dc offset voltages across the primary. In most practical transformer designs, normal tolerances on winding radius and wire resistance make this a less precise method than the resistor pair. Virtually all microphone input transformers will require loading on the secondary to control high-frequency response. For the circuit in the figure, network R_1, R_2, and C_1 shape the high-frequency response to a Bessel roll-off curve. Because they operate at very low signal levels, most microphone input transformers also include magnetic shielding.

11.2.1.2 Line Input

A line input transformer is driven by a balanced line and, most often, drives a ground-referenced (unbalanced) amplifier stage. As discussed in Chapter 37, modern voltage-matched interconnections require that line inputs have impedances of 10 kΩ or more, traditionally called *bridging* inputs. In the circuit of Fig. 11-30, a 4:1 step-down transformer is used which has an input impedance of about 40 kΩ.

High CMRR is achieved in line input transformers using the same techniques as those for microphones. Again, because its common-mode input impedances consist of small capacitances, a good input transformer will exhibit high CMRR even when signal sources are real-world equipment with less-than-perfect output impedance balance. The dirty little secret of most electronically balanced input stages, especially simple differential amplifiers, is that they are very susceptible to tiny impedance imbalances in driving sources. However, they usually have impressive CMRR figures when the driving source is a laboratory generator. The pitfalls of measurement techniques will be discussed in section 11.3.1.

As with any transformer having a Faraday shield, line input transformers have significant leakage inductance and their secondary load effectively controls their high-frequency response characteristics. The load resistance or network recommended by the manufacturer should be used to achieve specified bandwidth and transient response. Input transformers are intended to immediately precede an amplifier stage with minimal input capacitance. Additional capacitive loading of the secondary should be avoided because of its adverse effect on frequency and phase response. For example, capacitive loads in excess of about 100 pF—about 3 ft of standard shielded cable—can degrade performance of a standard 1:1 input transformer.

11.2.1.3 Moving-Coil Phono Input

Moving-coil phonograph pickups are very low-impedance, very low-output devices. Some of them have source impedances as low as 3 Ω, making it nearly impossible to achieve optimum noise performance in an amplifier. The transformer shown in Fig. 11-31 has a

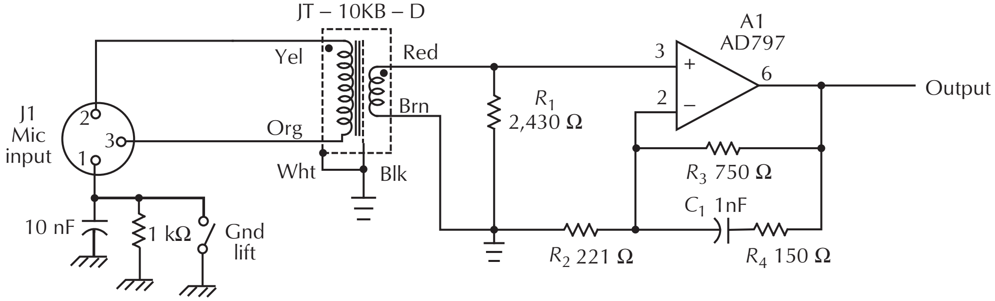


Figure 11-30. Low-noise unity-gain balanced line input stage.

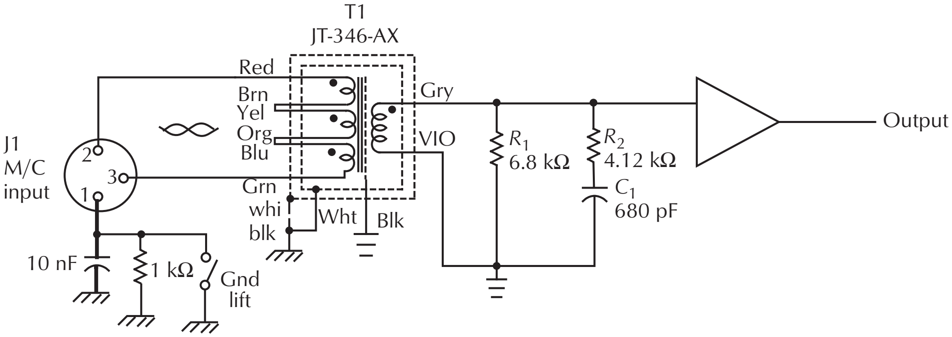


Figure 11-31. Preamp for 25 Ω moving-coil pickups.

three-section primary that can be series-connected as a 1:4 step-up for 25 Ω to 40 Ω devices and parallel-connected as a 1:12 step-up for 3 Ω to 5 Ω devices. In either case, the amplifier sees a 600 Ω source impedance that optimizes low-noise operation. The transformer is packaged in double magnetic shield cans and has a Faraday shield. The loading network R_1, R_2, and C_1 tailor the high-frequency response to a Bessel curve.

11.2.1.4 Line Output

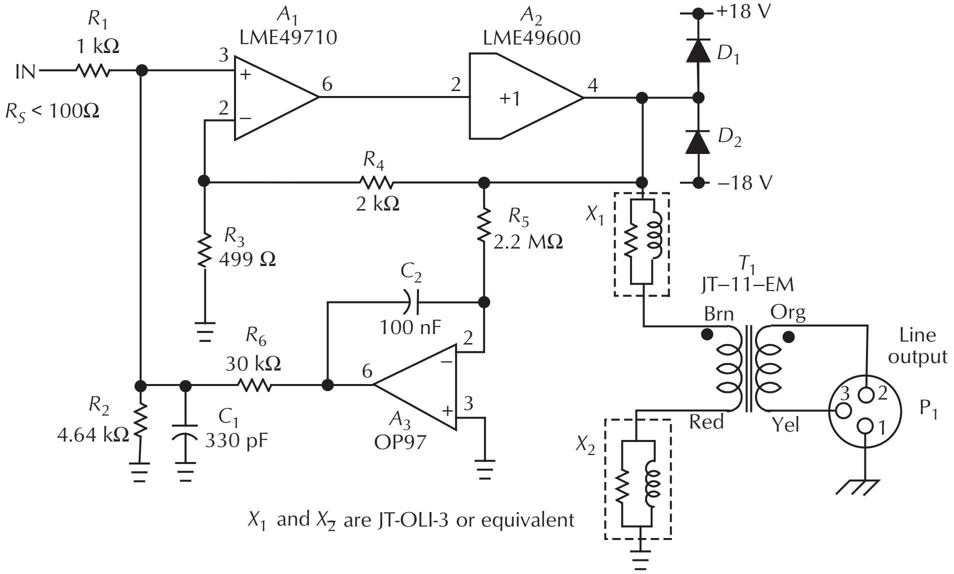
A line-level output transformer is driven by an amplifier and typically loaded by several thousand pF of cable capacitance plus the 20 kΩ input impedance of a balanced bridging line receiver. At high frequencies, most driver output current is actually used driving the cable capacitance. Sometimes, terminated 150 Ω or 600 Ω lines must be driven, requiring even more driver output current. Therefore, a line output transformer must have a low output impedance that stays low at high frequencies. This requires both low resistance windings and very low leakage inductance, since they are effectively in series between amplifier and load. To maintain impedance balance of the output line, both driving impedances and inter-winding capacitances must be well matched at each end of the windings. A typical bi-filar-wound design has winding resistances of 40 Ω each, leakage inductance of a few micro-henries, and a total inter-winding capacitance of about 20 nF matched to within 2% across the windings.

The high-performance circuit of Fig. 11-32 uses op-amp A_1 and current booster A_2 in a feedback loop setting overall gain at 12 dB. A_3 provides the high gain for a dc servo feedback loop used to keep dc offset at the output of A_2 under 100 µV. This prevents any significant dc flow in the primary of transformer T_1. X_1 provides capacitive load isolation for the amplifier and X_2 serves as a tracking impedance to maintain high-frequency impedance balance of the output. High-conductance diodes D_1 and D_2 clamp inductive kick to protect A_2 in case an unloaded output is driven into hard clipping.

The circuit of Fig. 11-33 is well suited to the lower signal levels generally used in consumer systems. Because its output floats, it can drive either balanced or unbalanced outputs, but not at the same time. Floating the unbalanced output avoids ground loop problems that are inherent to unbalanced interconnections.

In both previous circuits, because the primary drive of T_1 is single-ended, the voltages at the secondary will *not* be symmetrical, especially at high frequencies. THIS IS NOT A PROBLEM. Contrary to widespread myth and as explained in Chapter 37, signal symmetry has absolutely nothing to do with noise rejection in a balanced interface! Signal symmetry in this, or any other floating output, will depend on the magnitude and matching of cable and load impedances to ground. If there is a requirement for signal symmetry, the transformer should be driven by dual, polarity-inverted drivers.

The circuit of Fig. 11-34 uses a cathode follower circuit which replaces the usual resistor load in the cathode with an active current sink. The circuit operates at quiescent plate currents of about 10 mA and presents a driving source impedance of about 60 Ω to the transformer, which is less than 10% of its primary dc resistance. C_2 is used to prevent dc flow in the primary. Since the transformer has a 4:1 turns ratio, or 16:1 impedance ratio, a 600 Ω output load is reflected back to the driver circuit as about 10 kΩ. Since the signal swings on the primary are four times as large as those on the secondary, high-frequency capacitive coupling is prevented by a Faraday shield. The secondary windings may be parallel connected to drive a 150 Ω load. Because of the Faraday shield, output winding capacitances are low and the output signal


R1
1 kΩ
IN
R_S < 100Ω
A1
LME49710
3
+
2
−
6
A2
LME49600
2
+1
4
+18 V
D1
D2
−18 V
R4
2 kΩ
R3
499 Ω
R5
2.2 MΩ
X1
C2
100 nF
T1
JT–11–EM
Brn
Org
Line
output
R6
30 kΩ
6
2
−
+
3
A3
OP97
R2
4.64 kΩ
C1
330 pF
Red
Yel
X2
3
2
1
P1

X1 and X2 are JT-OLI-3 or equivalent

Figure 11-32. Typical line output application circuit.

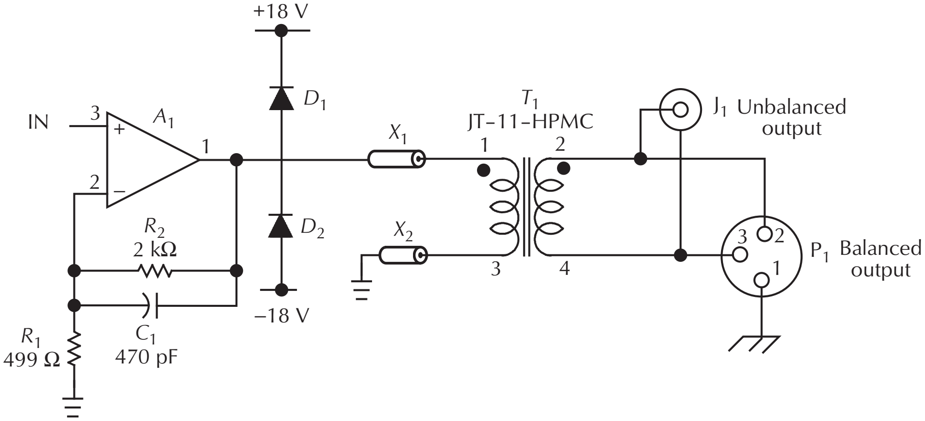

+18 V
IN
3
+
A1
2
−
1
D1
D2
−18 V
X1
X2
T1
JT–11–HPMC
1
2
3
4
J1 Unbalanced
output
3
2
1
P1 Balanced
output
R2
2 kΩ
C1
470 pF
R1
499 Ω

Figure 11-33. Universal isolated output application.

symmetry will be determined largely by the balance of line and load impedances.

11.2.1.5 Inter-Stage and Power Output

Inter-stage coupling transformers are seldom seen in contemporary equipment but were once quite popular in vacuum-tube amplifier designs. They typically use turns ratios in the 1:1 to 1:3 range and, as shown in Fig. 11-35, may use a center-tapped secondary producing phase-inverted signals to drive a push-pull output stage. Because both plate and grid circuits are relatively high impedance, windings are sometimes section-wound to reduce capacitances. Resistive loading of the secondary is usually necessary both to provide damping and to present a uniform load impedance to the driving stage. Although uncommon, inter-stage transformers for solid-state circuitry are most often bi-filar wound units similar to line output designs.

The classic push-pull *power output* stage, with many variations over the years, has been used in hi-fi gear, PA systems, and guitar amplifiers. The turns ratio of the output transformer is generally chosen for a reflected load at the tubes of several thousand ohms plate-to-plate. A typical 30:1 turns ratio may require

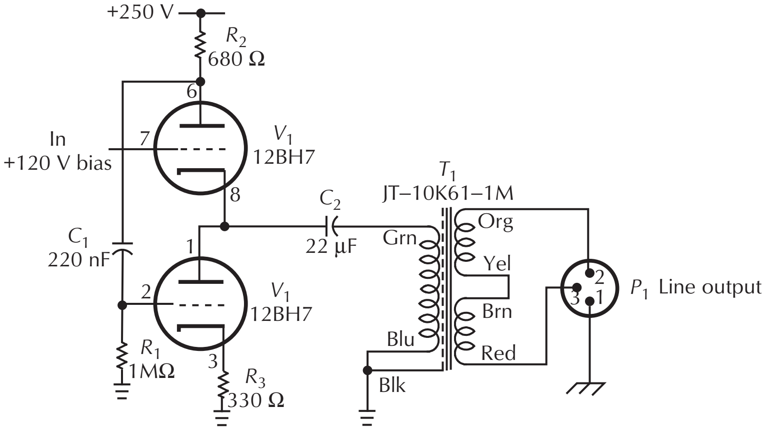

+250 V
R_2
680 Ω
6
In
+120 V bias
7
V_1
12BH7
8
C_2
22 µF
T_1
JT–10K61–1M
Org
Grn
Yel
Brn
Blu
Red
Blk
C_1
220 nF
1
2
V_1
12BH7
R_1
1MΩ
3
R_3
330 Ω
2
3
1
P_1 Line output

Figure 11-34. Double cathode-follower line driver.

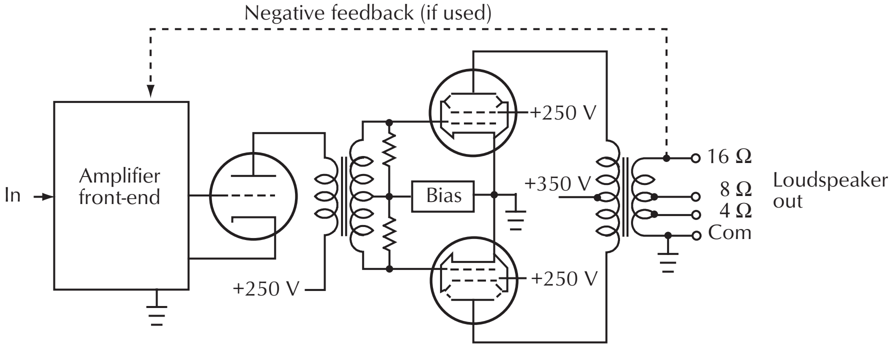

Negative feedback (if used)
In
Amplifier front-end
+250 V
Bias
+250 V
+350 V
+250 V
16 Ω
8 Ω
4 Ω
Com
Loudspeaker out

Figure 11-35. Push-pull vacuum-tube power amplifier.

many interleaved sections to achieve bandwidth extending well beyond 20 kHz.

If the quiescent plate currents and the number of turns in each half of the primary winding are matched, magnetic flux in the core will cancel at dc. Since any current-balancing in vacuum-tubes is temporary at best, these transformers nearly always use steel cores to better tolerate this unbalanced dc in their windings. The relatively high driving impedance of the tube plates results in considerable transformer related distortion. To reduce distortion, feedback around the transformer is often employed. To achieve stability (freedom from oscillation), very wide bandwidth (actually low phase shift) is required of the transformer when a feedback loop is closed around it. As a result, some of these output transformer designs are very sophisticated. Some legendary wisdom suggests as a rough guide that a good-fidelity output transformer should have a core weight and volume of at least 0.34 pounds and 1.4 cubic inches respectively per watt of rated power.[4]

A single-ended power amplifier is created by removing the lower tube and the lower half of the transformer primary from the circuit of Fig. 11-35. Now plate current will create a strong dc field in the core. As discussed in section 11.1.2.1, the core will likely require an air gap to avoid magnetic saturation. The air gap reduces inductance, limiting low-frequency response, and increases even-order distortion products. Such a single-ended pentode power amplifier was ubiquitous in classic AM 5-tube table radios of the fifties and sixties.

11.2.1.6 Microphone Output

There are two basic types of output transformers used in microphones, step-up and step-down. In a ribbon microphone, the ribbon element may have an impedance of well under 1 Ω, requiring a step-up transformer with a turns ratio of 1:12 or more to raise its output level and make its nominal output impedance around 150 Ω. Typical dynamic elements have impedances from 10 Ω to 30 Ω, which require step-up turns ratios from 1:2 to 1:4. These step-up designs are similar to line output transformers in that they have no Faraday or magnetic shields, but are smaller because they operate at lower signal levels.

A condenser microphone has integral circuitry to buffer and/or amplify the signal from its extremely high-impedance transducer. Since this low-power circuitry operates from the phantom supply, it may be unable to directly drive the 1.5 kΩ input impedance of a typical microphone preamp. The output transformer shown in Fig. 11-36, which has an 8:1 step-down ratio, will increase the impedance seen by Q_1 to about 100 kΩ. Because of its high turns ratio, a Faraday shield is used to prevent capacitive coupling of the primary signal to the output.

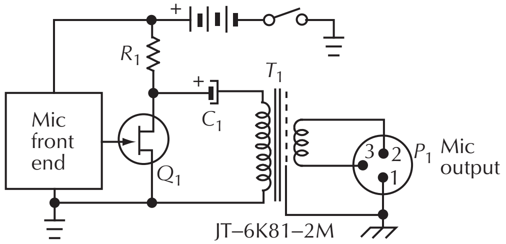


Figure 11-36. Condenser microphone output transformer.

11.2.2 System-Level Applications

11.2.2.1 Microphone Isolation or Splitter

The primary of a transformer with a 1:1 turns ratio can bridge the output of a 150 Ω to 200 Ω microphone feeding one pre-amp and the secondary of the transformer can feed a duplicate of the microphone signal to another pre-amp. Of course, a simple Y cable could do this but there are potential problems. There are often large and noisy voltages between the grounds of two pre-amplifiers. The isolation provided by the transformer prevents the noise from coupling to the balanced signal line. To reduce capacitive noise coupling, Faraday shields are included in better designs and double Faraday shields in the best. As discussed in Section 11.1.3.5, the input impedances of all the pre-amps, as well as all the cable capacitances, will be seen in parallel by the microphone. This places a practical upper limit on how many ways the signal can be split. Transformers are commercially available in 2, 3, and 4-winding versions. A 3-way splitter box schematic is shown in Fig. 11-37. Since the microphone is *directly* connected only to the direct output, it is the only one that can pass phantom power to the microphone. To each preamp, each isolated output looks like a normal floating (ungrounded) microphone. The ground lift switches are normally left open to prevent possible high ground current flow in the cable shields.

11.2.2.2 Microphone Impedance Conversion

Some legacy dynamic microphones are high-impedance, about 50 kΩ, and have two-conductor cable and connector (unbalanced). When such a microphone must be connected to a standard balanced low-impedance microphone pre-amp, a transformer with a turns ratio of about 15:1 is necessary. Similar transformers can be used to adapt a low-impedance microphone to the unbalanced high-impedance input of a legacy pre-amplifier. Commercial products are available which enclose such a transformer in the XLR adapter barrel.

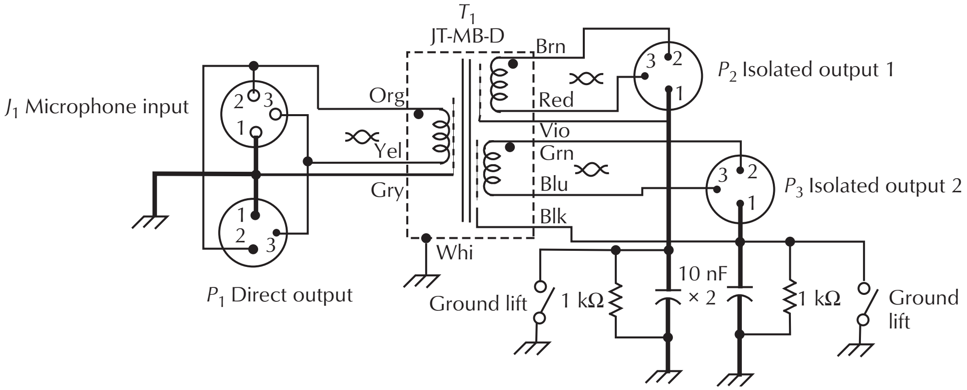


Figure 11-37. A 3-way microphone splitter box.

11.2.2.3 Line to Microphone Input or Direct Box

Because its high-impedance, unbalanced input accepts line-level signals and its output drives the low-level, low-impedance balanced microphone input of a mixing console, the device shown in Fig. 11-38 is called a direct box. It is most often driven by an electric guitar, synthesizer, or other stage instrument. Because it uses a transformer, it provides ground isolation as well. In this typical circuit, since the transformer has a 12:1 turns ratio, the impedance ratio is 144:1. When the microphone input has a typical 1.5 kΩ input impedance, the input impedance of the direct box is about 200 kΩ. The transformer shown has a separate Faraday shield for each winding to minimize capacitively coupled ground noise.

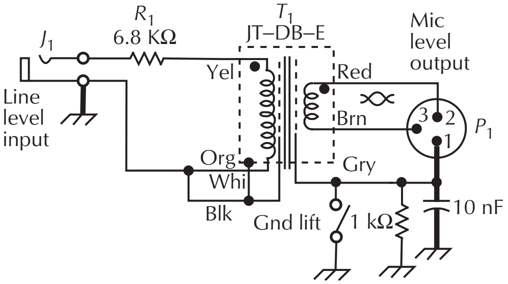


Figure 11-38. A transformer-isolated direct box.

11.2.2.4 Line Isolation or Hum Eliminators

There are a remarkable number of black boxes on the market intended to solve ground loop problems. This includes quite a number of transformer-based boxes. With rare exception, those boxes contain *output* transformers. Tests were performed to compare noise rejection of the original interface to one with an added output transformer and to one with an added input transformer. The tests accurately simulated typical real-world equipment, see the definitions at the end of this section.

Fig. 11-39 shows results of CMRR tests on a ***balanced*** interface using the IEC 60268-3 test procedure discussed in Section 11.3.1.2. This test recognizes that the impedances of real-world balanced outputs are not matched with the precision of laboratory equipment. While the output transformer reduces 60 Hz hum by over 20 dB, it has little effect on buzz artifacts over about 1 kHz. The input transformer increases rejection to over 120 dB at 60 Hz and to almost 90 dB at 3 kHz, where the human ear is most sensitive to faint sounds.

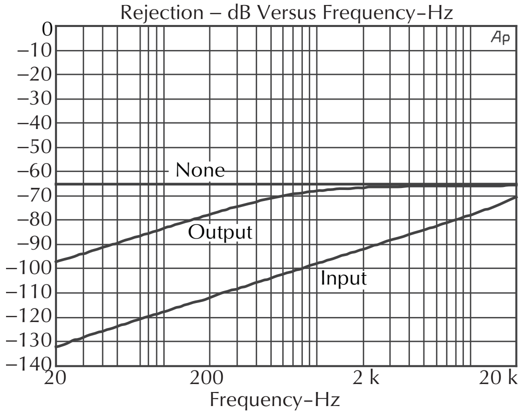


Figure 11-39. Balanced output to balanced input.

Fig. 11-40 shows results of ground noise rejection tests on an ***unbalanced*** interface. By definition, there is 0 dB of inherent rejection in an unbalanced interface, see Chapter 37.While the output transformer reduces 60 Hz hum by about 70 dB, it reduces buzz artifacts around 3 kHz by only 35 dB. The input transformer increases rejection to over 100 dB at 60 Hz and to over 65 dB at 3 kHz.

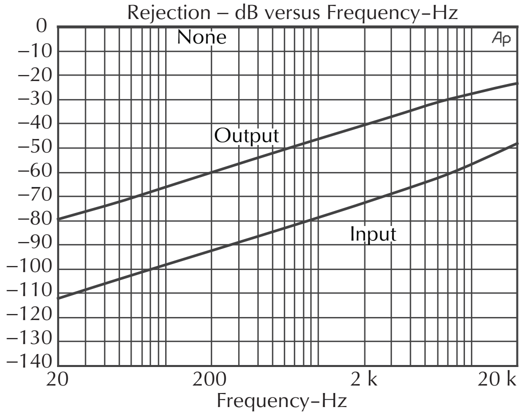


Figure 11-40. Unbalanced output to unbalanced input.

Fig. 11-41 shows results of CMRR tests when an ***unbalanced*** output drives a ***balanced*** input. A two-wire connection of this interface will result in zero rejection, see Chapter 37. Assuming a three-wire connection, the −30 dB plot shows how CMRR of typical electronically-balanced input stages is degraded by the 600 Ω source imbalance. Again, the output transformer improves 60 Hz hum by over 20 dB, it has little effect on buzz artifacts over about 1 kHz. The input transformer increases rejection to almost 100 dB at 60 Hz and to about 65 dB at 3 kHz.

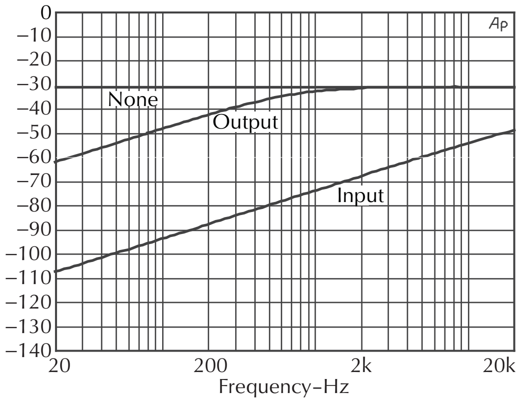


Figure 11-41. Unbalanced output to balanced input.

Fig. 11-42 shows results of ground noise rejection tests when a ***balanced*** output drives an ***unbalanced*** input. Because our balanced output does not float, the direct connection becomes an unbalanced interface having, by definition, 0 dB of rejection. While the output transformer reduces 60 Hz hum by about 50 dB, it reduces buzz artifacts around 3 kHz by less than 20 dB. The input transformer increases rejection to almost 100 dB at 60 Hz and to about 65 dB at 3 kHz. In this application it is usually desirable to attenuate the signal by about 12 dB—from +4 dBu or 1.228 V to −10 dBV or 0.316 V—as well as provide ground isolation. This can be conveniently done by using a 4:1 step-down input transformer such as the one in Fig. 11-29, which will produce rejection comparable to that shown here.

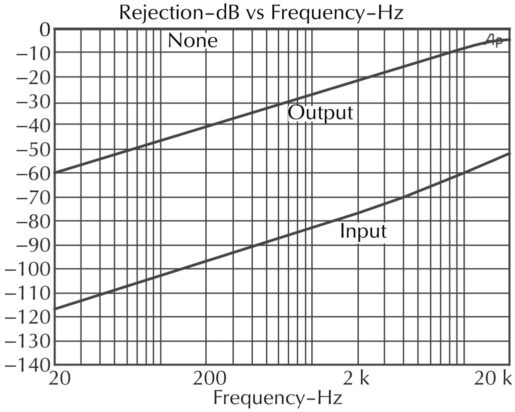


Figure 11-42. Balanced output to unbalanced input.

One might fairly ask “Why not use a 1:4 step-up transformer when an ***unbalanced*** output drives a ***balanced*** input to get 12 dB of signal gain?” Because of the circuit impedances involved, the answer is because it doesn’t work very well. Recall that a 1:4 turns ratio has an impedance ratio of 1:16. This means that the input impedance of the pro balanced input we drive will be reflected back to the consumer output at one-sixteenth that. Since the source impedance—usually unspecified, but not the same as *load* impedance—of a consumer outputs is commonly 1 kΩ or more, the reflected loading losses are high. A 1:4 step-up transformer would have its own insertion losses, which we will rather optimistically assume at 1 dB. The table below shows actual gain using this transformer with some typical equipment output and input impedances (*Z* is impedance).

Table 11-1. Gain Derived from a 1:4 Step-up Transformer in Typical Circuits

Consumer Output Z	Pro Balanced Input Z		
	10 kΩ **(625 Ω)**	**20 kΩ** **(1.25 kΩ)**	**40 kΩ** **(2.5 kΩ)**
200 Ω	8.6 dB	9.7 dB	10.3 dB
500 Ω	5.9 dB	8.1 dB	9.4 dB
1 kΩ	2.7 dB	5.9 dB	8.1 dB

Not only will gain usually be much less than 12 dB, the load reflected to the consumer output, shown in parentheses, is excessive and will likely cause high distortion, loss of headroom, and poor low-frequency response. Often the only specification of a consumer output is 10 kΩ minimum load. It is futile to increase the turns ratio of the transformer in an attempt to overcome the gain problem—it only makes the reflected loading problems worse. In most situations, a 1:1 transformer can be used because the pro equipment can easily provide the required gain. Of course, a 1:1 *input* transformer will provide far superior noise immunity from ground loops as well.

The point here is that the noise rejection provided by an input transformer with a Faraday-shield is far superior to that provided by an output type. But the input transformer must be used at the receiver or destination end of an interface cable. In general, input transformers should drive no more than three feet of typical shielded cable—the capacitance of longer cables will erode their high-frequency bandwidth. Although output type transformers without a Faraday shield are not as good at reducing noise, their advantage is that they can be placed anywhere along an interface cable, at the driver end, at a patch-bay, or at the destination end, and work equally well (or poorly, compared to an input transformer). In all the test cases discussed in this section, results of using *both* an output and an input transformer produced results *identical* to those using only an input transformer. For example, an unbalanced output does not need to be balanced by a transformer before transmission through a cable (this is a corollary of the balance versus symmetry myth), it needs only an input transformer at the receiver. There is rarely a need to ever use both types on the same line.

Definitions (in context of comparison tests only):

Balanced Output. ”A normal, non-floating source having a differential output impedance of 600 Ω and common-mode output impedances of 300 Ω, matched to within ± 0.1%.

Balanced Input. A typical electronically-balanced stage—an instrumentation circuit using 3 op-amps—having a differential input impedance of 40 kΩ and common-mode input impedances of 20 kΩ, trimmed for a CMRR over 90 dB when directly driven by the above Balanced Output.

Unbalanced Output. A ground-referenced output having an output impedance of 600 Ω. This is representative of typical consumer equipment.

Unbalanced Input. A ground-referenced input having an input impedance of 50 kΩ. This is representative of typical consumer equipment.

No Transformer. A direct wired connection.

Output Transformer. A Jensen JT-11-EMCF—a popular 1:1 line output transformer.

Input Transformer. A Jensen JT-11P-1—the most popular 1:1 line input transformer.

11.2.2.5 Loudspeaker Distribution or Constant Voltage

When a number of low-impedance loudspeakers are located far from a power amplifier, there are no good methods to interconnect them in a way that properly loads the amplifier. The problem is compounded by the fact that power losses due to the resistance of the inter-connecting wiring can be substantial. The wire gauge required is largely determined by the current it must carry and its length. Borrowing a technique from power utility companies, boosting the distribution voltage reduces the current for a given amount of power and allows smaller wire to be used in the distribution system. *Step-down* matching transformers, most often having taps to select power level and/or loudspeaker impedance, are used at each location. This scheme not only reduces the cost of wiring but allows system designers the freedom to choose how power is allocated among the speakers. These so-called constant-voltage loudspeaker distribution systems are widely used in public address, paging, and background music systems. Although the most popular is 70 V, others include 25 V, 100 V, and 140 V. Because the higher voltage systems offer the lowest distribution losses for a given wire size, they are more common in very large systems. It should also be noted that only the 25 V system is considered low-voltage by most regulatory agencies and the wiring in higher voltage systems may need to conform to power wiring practices.

It is important to understand that these nominal voltages exist on the distribution line only when the driving amplifier is operating at full rated power. Many specialty power amplifiers have outputs rated to drive these lines directly but ordinary power amplifiers rated to drive speakers can also drive such lines, according to Table 11-2.

Table 11-2. Amplifier Power Required at Various Impedances Versus Output Voltage

Amplifier Rated Output, Watts			Output Voltage
at 8 Ω	at 4 Ω	at 2 Ω	
1250	2500	5000	100
625	1250	2500	70.7
312	625	1250	50
156	312	625	35.3
78	156	312	25

For example, an amplifier rated to deliver 1,250 W of continuous average power into an 8 Ω load will drive a 70 V distribution line directly as long as the sum of the power delivered to all the loudspeakers doesn’t exceed 1,250 W. Although widely used, the term *rms watts* is technically ambiguous.[5] In many cases, the benefits of constant-voltage distribution are desired, but the total power required is much less. In that case a *step-up* transformer can be used to increase the output voltage of an amplifier with less output. This is often called matching it to the line because such a transformer is actually transforming the equivalent line impedance down to the rated load impedance for the amplifier. Most of these step-up transformers will have a low turns ratio. For example, a 1:1.4 turns ratio would increase the 50 V output to 70 V for an amplifier rated at 300 W into 8 Ω. In such low-ratio applications, the *auto-transformer* discussed in Section 11.1.2.2 has cost and size advantages. Fig. 11-43 is a schematic of an auto-transformer with taps for turns ratios of 1:1.4 or 1:2 which could be used to drive a 70 V line from amplifiers rated for either 300 or 150 W respectively at 8 Ω. Several power amplifier manufacturers offer such transformers as options or accessories.

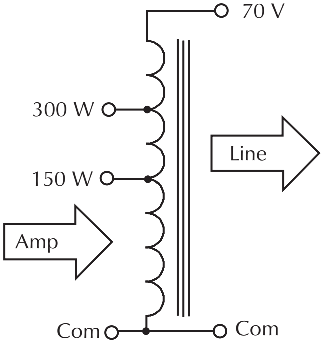


Figure 11-43. Step-up auto-transformer

A line to voice-coil transformer is usually necessary to step-down the line voltage and produce the desired loudspeaker power, Table 11-3.

These step-down transformers can be designed several ways. Fig. 11-44 shows a design where the line voltage is selected at the primary side and the power level is selected at the secondary while Fig. 11-45 shows a design where power level is selected on the primary side and loudspeaker impedance is selected at the secondary.

As may be seen from the repeating patterns in the table above, there are many combinations of line voltage, loudspeaker impedance, and power level that result in the same required turns ratio in the matching transformer.

Since the constant-voltage line has a very low source impedance, and the transformer is loaded by a low-impedance loudspeaker, transformer high-frequency response is usually not a design issue. As in any transformer, low-frequency response is determined by primary inductance and total source impedance, which is dominated by the primary winding resistance since the driving source impedance is very low. Winding resistances of both primary and secondary contribute to insertion loss. In efforts to reduce size and cost, the fewest turns of the smallest wire possible are often used, which raises insertion loss and degrades low-frequency response. Generally, an insertion loss of 1 dB or less is considered good and 2 dB is marginally acceptable for these applications.

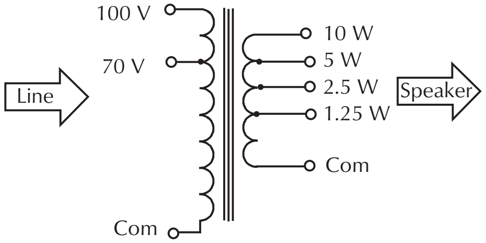


Figure 11-44. Transformer with secondary taps for power selection.

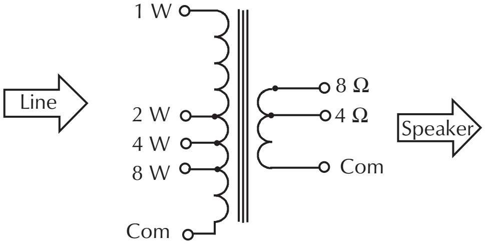


Figure 11-45. Transformer with primary taps for power selection.

Table 11-3. Transformer Step-Down Turns Ratio Required to Produce the Desired Loudspeaker Power

Speaker Power in Watts			Loudspeaker	Transformer Step-Down Turns Ratio Required			
16 Ω	8 Ω	4 Ω	Volts	100 V	70 V	35 V	25 V
32	64	128	22.63	4.42	3.12	1.56	1.10
16	32	64	16.00	6.25	4.42	2.21	1.56
8	16	32	11.31	8.84	6.25	3.12	2.21
4	8	16	8.00	12.50	8.84	4.42	3.12
2	4	8	5.66	17.70	12.50	6.25	4.42
1	2	4	4.00	25.00	17.70	8.84	6.25
0.5	1	2	2.83	35.30	25.00	12.50	8.84
0.25	0.5	1	2.00	50.00	35.30	17.70	12.50
0.125	0.25	0.5	1.41	71.00	50.00	25.00	17.70

It is very important to understand that, while the low-level frequency response of a transformer may be rated as −1 dB at 40 Hz, its rated power does NOT apply at that frequency. Rated power, or maximum signal level is discussed in Section 11.1.3.1. In general, level handling is increased by more primary turns and more core material and it takes more of both to handle more power at lower frequencies. This ultimately results in physically larger, heavier, and more expensive transformers. When any transformer is driven at its rated level at a lower frequency than its design will support, magnetic core saturation is the result. The sudden drop in permeability of the core effectively reduces primary inductance to zero. The transformer primary now appears to have only the dc resistance of its winding, which may be only a fraction of an ohms. In the best scenario, some ugly-sounding distortion will occur and the line amplifier will simply current limit. In the worst scenario, the amplifier will not survive the inductive energy fed back as the transformer comes out of saturation. This can be especially dangerous if large numbers of transformers saturate simultaneously.

In 1953, the power ratings of loudspeaker matching transformers were based on 2% distortion at 100 Hz.[6] Traditionally, the normal application of these transformers has been speech systems and this power rating standard assumes very little energy will exist under 100 Hz. The same reference recommends that transformers used in systems with emphasized bass should have ratings higher than this 100 Hz nominal power rating and those used to handle organ music should have ratings of at least four times nominal. Since the power ratings for these transformers is rarely qualified by an honest specification stating the applicable frequency, it seems prudent to assume that ***the historical 100 Hz power rating applies to most commercial transformers***.

If a background music system, for example, requires good bass response, it is wise to use over-rated transformers. Reducing the voltage on the primary side of the transformer will extends its low-frequency power handling. Its possible, using the table above, to use different taps to achieve the same ratio while driving less than nominal voltage into the transformer primary. For example, a 70 V line could be connected to the 100 V input of the transformer in Fig. 11-33 and, for example, the 10 W secondary tap used to actually deliver 5 W. In any constant-voltage system, saturation problems can be reduced by appropriate high-pass filtering. Simply attenuate low-frequency signals before they can reach the transformers. In voice-only systems, problems that arise from breath pops, dropped microphones, or signal switching transients can be effectively eliminated by a 100 Hz high-pass filter ahead of the power amplifier. In music systems, attenuating frequencies too low for the speakers to reproduce can be similarly helpful.

11.2.2.6 Telephone Isolation or Repeat Coil

In telephone systems it was sometimes necessary to isolate a circuit which was grounded at both ends. This metallic circuit problem was corrected with a repeat coil to improve longitudinal balance. Translating from telephone lingo, this balanced line had poor common-mode noise rejection which was corrected with a 1:1 audio isolation transformer. The Western Electric 111C repeat coil was widely used by radio networks and others for high-quality audio transmission over 600 Ω phone lines. It has split primary and secondary windings and a Faraday shield. Its frequency response was 30 Hz to 15 kHz and it had less than 0.5 dB insertion loss. Split windings allow them to be parallel connected for 150 Ω use.

Fig. 11-46 shows a modern version of this transformer as a general purpose isolator for low-impedance circuits, such as in a recording studio patch-bay. Optional components can be useful in some applications. For example, network R_1 and C_1 will flatten the input impedance over frequency, R_2 will trim the input impedance to exactly 600 Ω, and R_3 can be used to properly load the transformer when the external load is high-impedance or bridging.

11.2.2.7 Telephone Directional Coupling or Hybrid

Telephone hybrid circuits use bridge-nulling principles to separate signals which may be transmitted and received simultaneously or full-duplex on a 2-wire line. This nulling depends critically on well-controlled impedances in all branches of the circuits. This nulling is what suppresses the transmit signal (your own voice) in the receiver of your phone while allowing you to hear the receive signal (the other party).

A two-transformer hybrid network is shown in Fig. 11-47. The arrows and dashed lines show the current flow for a signal from the transmitter TX. Remember that the dots on the transformers show points having the same instantaneous polarity. The transformer turns ratios are assumed to be 1:1:1. When balancing network Z_N has an impedance that matches the line impedance Z_L at all significant frequencies, the currents in the Z_L loop (upper) and Z_N loop (lower) will be equal. Since they flow in opposite directions in the RX transformer (right), there is cancellation and the TX signal does not appear at RX. A signal originating from the line rather

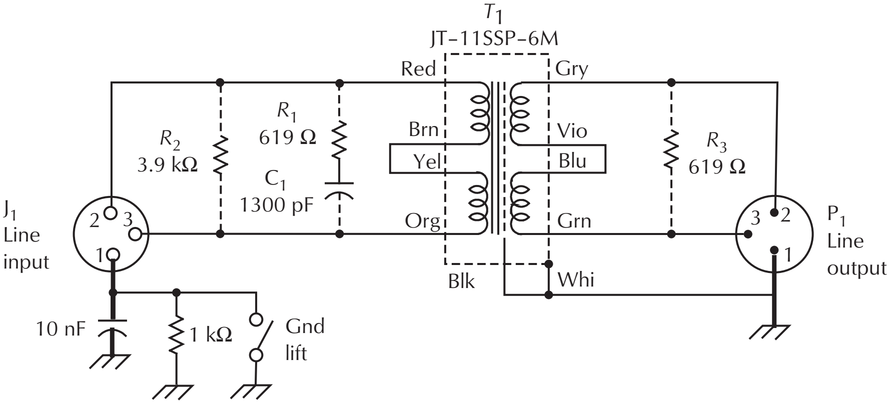

T_1
JT–11SSP–6M
Red
Gry
R_2
3.9 kΩ
R_1
619 Ω
C_1
1300 pF
Brn
Yel
Org
Vio
Blu
Grn
R_3
619 Ω
J_1
Line
input
2
3
1
3
2
1
P_1
Line
output
Blk
Whi
10 nF
1 kΩ
Gnd
lift

Figure 11-46. Repeat coil ground isolation for 600 Ω lines.

than TX is not suppressed and is heard in RX. A common problem with hybrids of any kind is adjusting network Z_N to match the telephone line, which may vary considerably in impedance even over relatively short time spans.

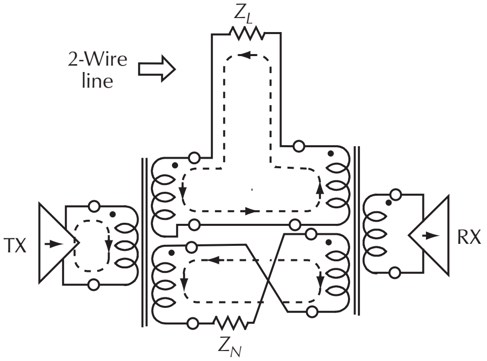

Z_L
2-Wire
line
TX
RX
Z_N

Figure 11-47. Two-transformer hybrid.

If the transmitter and receiver are electrically connected, the single transformer method, shown in Fig. 11-48, can be used. Any well-designed transformers with accurate turns ratio can be used in hybrid applications.

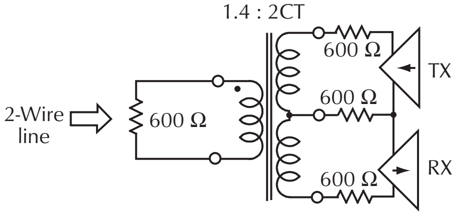

1.4 : 2CT
600 Ω
TX
600 Ω
2-Wire
line
600 Ω
600 Ω
RX

Figure 11-48. Single-transformer hybrid.

11.2.2.8 Moving-Coil Phono Step-Up

Outboard boxes are sometimes used to adapt the output of low-output, low-impedance moving-coil phono pickups to pre-amplifier inputs intended for conventional high-impedance moving-magnet pickups. These pre-amplifiers have a standard input impedance of 47 kΩ. Fig. 11-49 shows a 1:37 step-up transformer used for this purpose. It has a voltage gain of 31 dB and reflects its 47 kΩ pre-amplifier load to the pickup as about 35 Ω. This keeps loading loss on a 3 Ω pickup to about 1 dB. The series RC network on the secondary provides proper damping for smooth frequency response. Double magnetic shield cans are used because of the very low signal levels involved and the low-frequency gain inherent in the RIAA playback equalization. In these applications, it is extremely important to keep all leads to the pickup tightly twisted to avoid hum from ambient magnetic fields.

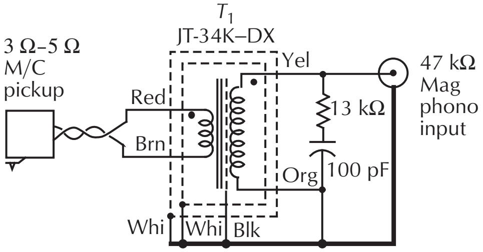

T_1
JT-34K–DX
3 Ω–5 Ω
M/C
pickup
Yel
47 kΩ
Mag
phono
input
Red
13 kΩ
Brn
Org
100 pF
Whi
Whi
Blk

Figure 11-49. Step-up transformer for moving coil phono pickup.

11.3 Measurements and Data Sheets

11.3.1 Testing and Measurements

11.3.1.1 Transmission Characteristics

The test circuits below are the basic setups to determine the signal transmission characteristics of output and input type transformers, respectively, shown in the diagrams as DUT for device under test. In each case, the driving source impedance must be specified and is split into two equal parts for transformers specified for use in balanced systems. For example, if a 600 Ω balanced source is specified, the resistors $R_s/2$ become 300 Ω each. The generator indicated in both diagrams is understood to have symmetrical voltage outputs. The buffer amplifiers shown are used to provide a zero source impedance, which is not available from most commercial signal sources. The generator could be used in an unbalanced mode by simply connecting the lower end of the DUT primary to ground. The specified load impedance must also be placed on the secondary. For output transformers, the load and meter are often floating as shown in Fig. 11-50. For input transformers, a specified end of the secondary is generally grounded as shown in Fig. 11-51.

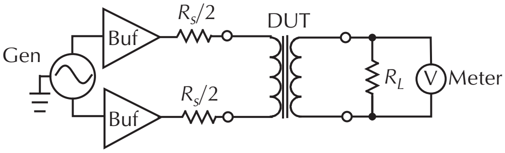


Figure 11-50. Transmission tests for output types.

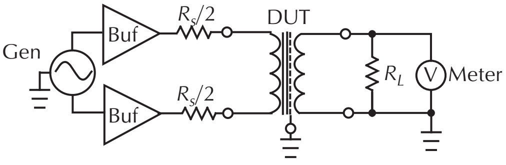


Figure 11-51. Transmission tests for input types.

These test circuits can be used to determine voltage gain or loss, turns ratio when R_L is infinite, frequency response, and phase response. If the meter is replaced with a distortion analyzer, distortion and maximum operating level may be characterized. Multi-purpose equipment such as the Audio Precision System 1 or System 2 can make such tests fast and convenient. Testing of high-power transformers usually requires an external power amplifier to boost the generator output as well as some hefty power resistors to serve as loads.

11.3.1.2 Balance Characteristics

Tests for common-mode rejection are intended to apply a common-mode voltage through some specified resistances to the transformer under test. Any differential voltage developed then represents undesired conversion of common-mode voltage to differential mode by the transformer. In general terms, CMRR or common-mode rejection ratio, is the ratio of the response of a circuit to a voltage applied normally (differentially) to that of the same voltage applied in common-mode through specified impedances. This conversion is generally the result of mismatched internal capacitances in the balanced winding. For output transformers, the most common test arrangement is shown in Fig. 11-52. Common values are 300 Ω for RG and values from zero to 300 Ω for $R_s/2$. Resistor pairs must be very well matched.

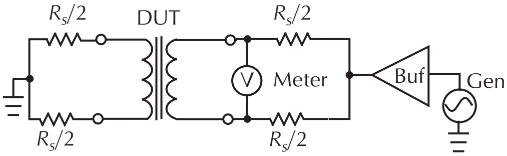


Figure 11-52. Common-mode test for output types.

Traditionally, CMRR tests of balanced input stages involved applying the common-mode voltage through a pair of very tightly-matched resistors. As a result, such traditional tests were not accurate predictors of real-world noise rejection for the overwhelming majority of electronically-balanced inputs. The IEC recognized this in 1998 and solicited suggestions to revise the test. The problem arises from the fact that the common-mode output impedances of balanced sources in typical commercial equipment are not matched with laboratory precision. Imbalances of 10 Ω are quite common. This author, through an educational process about balanced interfaces in general, suggested a more realistic test which was eventually adopted by the IEC in their standards document 60268-3 "Testing of Amplifiers" in August, 2000. The "Informative Annex" of this document is a concise summary explaining the nature of a balanced interface. The method of the new test, as shown in Fig. 11-53, is simply to introduce a 10 Ω imbalance, first in one line and then in the other. The CMRR is then computed based on the highest differential reading observed.

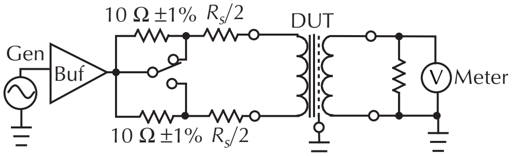


Figure 11-53. IEC Common-mode test for input types.

11.3.1.3 Resistances, Capacitances, and Other Data

Other data which can be very helpful to an equipment or system designer includes resistances of each winding and capacitances from winding to winding or winding to Faraday shield or transformer frame. Do not use an ohmmeter to check winding resistances unless you are able to later demagnetize the part. Ordinary ohmmeters, especially on low-ohm ranges, can weakly magnetize the core. If an ohmmeter simply must be used, use the highest range where the current is least.

Capacitances are usually measured on impedance bridges and, to minimize the effects of winding inductances, with all windings shorted. Total capacitances can be measured this way, but balance of capacitances across a winding must be measured indirectly. CMRR tests are effectively measuring capacitance imbalances.

As shown in Fig. 11-54, sometimes the input impedance of a winding is measured with specified load on other windings. This test includes the effects of primary resistance, secondary resistance, and the parallel loss resistance RC shown in Fig. 11-8 and Fig. 11-13. If specified over a wide frequency range, it also includes the effects of primary inductance and winding capacitances.

Breakdown voltages are sometimes listed as measures of insulation integrity. This is normally done with special equipment, sometimes called a hi-pot tester, which applies a non-destructive high voltage while limiting current to a very low value.

11.3.2 Data Sheets

11.3.2.1 Data to Impress or to Inform?

Data sheets and specifications exist to allow easy comparison of one product with others. But, in a world where marketing seems to supersede all else, honest data sheets and guaranteed specifications are becoming increasingly rare. As with many other audio products, most so-called data sheets and specifications are designed to impress rather than inform. **Specifications offered with unstated measurement conditions are essentially meaningless**, so a degree of skepticism is always appropriate before comparisons are made. A few examples:

- Hum Eliminator and Line Level Shifter products with no noise rejection or CMRR specs at all!
- Line Level Shifter products with no gain spec at all! Section 11.2.2.4 explains why they won't tell you!
- Maximum Power or Maximum Level listed with no frequency and no source impedance specified!

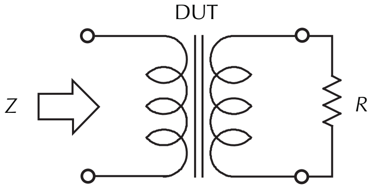


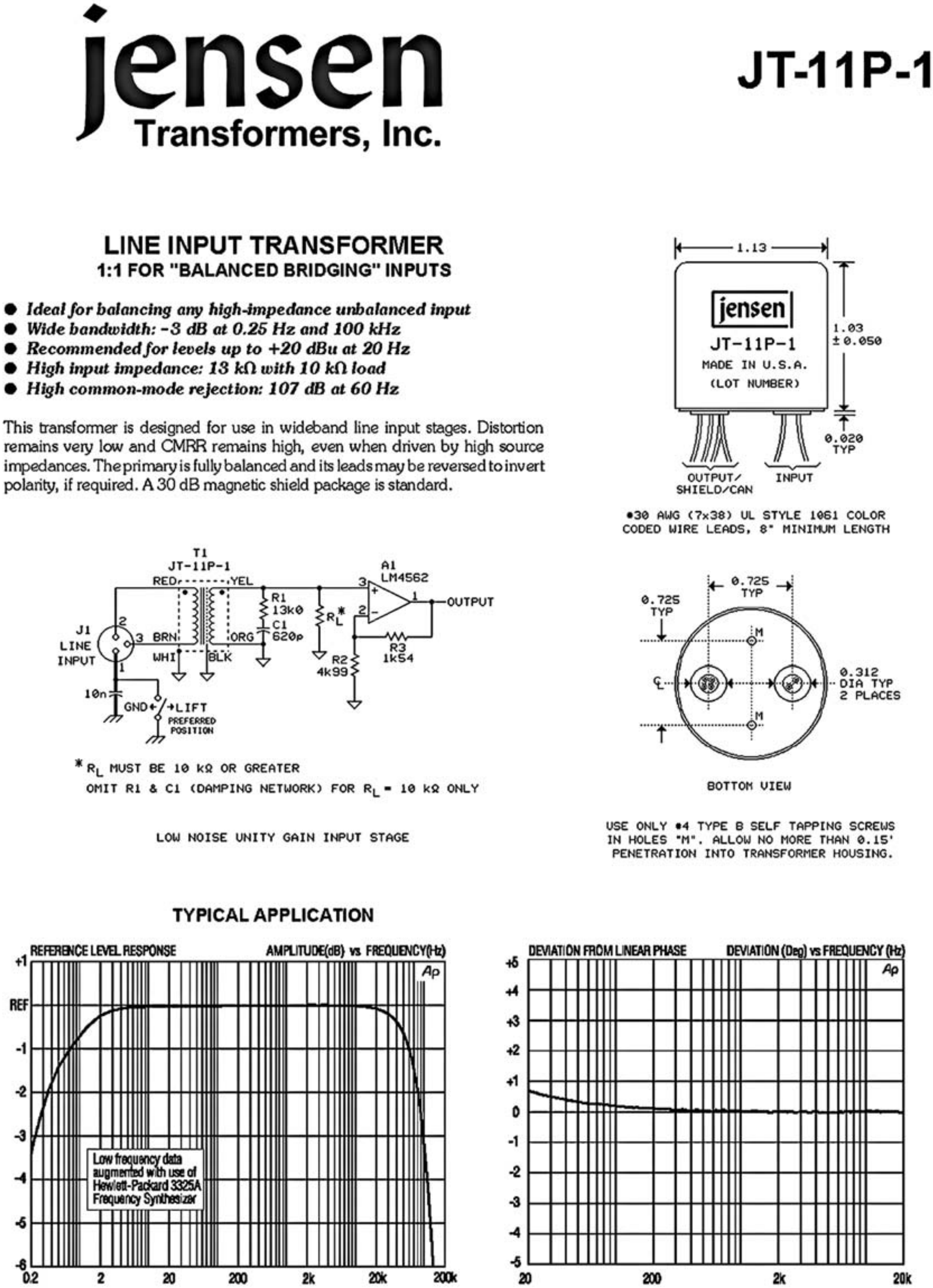
Figure 11-54. Impedance tests.

Other specifications, while technically true, are likely to mislead those not wise in the ways of transformers. For example, Maximum Level and Distortion are commonly specified at 50 Hz, 40 Hz, or 30 Hz instead of the more rigorous 20 Hz. Be careful, specs at these higher frequencies will always be much more impressive than those at 20 Hz! There is an approximate 6 dB per octave relationship at work here. A transformer specified for level or distortion at 40 Hz, for example, will handle about 6 dB less level at 20 Hz and have *at least* twice the distortion!

Seen in transformer-based hum eliminator advertising copy: "Frequency response 10 Hz to 40 kHz ±1 dB into 10 kΩ load" and "Distortion less than 0.002% at 1 kHz." What about the source impedance? Response at 10 Hz and distortion are always much better when a transformer is driven from a 0 Ω source! What happens when a real-world source drives the box? For a full-range audio transformer, measuring distortion at 1 kHz is nearly meaningless. Section 11.1.3.1 explains.

11.3.2.2 Comprehensive Data Sheet Example

For reference, the following is offered as a sample of a data sheet that has been called truly useful and brutally honest. Note that minimum or maximum limits are guaranteed for the most critical specifications!



jensen
Transformers, Inc.

JT-11P-1

LINE INPUT TRANSFORMER
1:1 FOR "BALANCED BRIDGING" INPUTS

- *Ideal for balancing any high-impedance unbalanced input*
- *Wide bandwidth: –3 dB at 0.25 Hz and 100 kHz*
- *Recommended for levels up to +20 dBu at 20 Hz*
- *High input impedance: 13 kΩ with 10 kΩ load*
- *High common-mode rejection: 107 dB at 60 Hz*

This transformer is designed for use in wideband line input stages. Distortion remains very low and CMRR remains high, even when driven by high source impedances. The primary is fully balanced and its leads may be reversed to invert polarity, if required. A 30 dB magnetic shield package is standard.

1.13
jensen
JT-11P-1
MADE IN U.S.A.
(LOT NUMBER)
1.03 ± 0.050
0.020 TYP
OUTPUT/ SHIELD/CAN
INPUT

•30 AWG (7x38) UL STYLE 1061 COLOR CODED WIRE LEADS, 8" MINIMUM LENGTH

T1
JT-11P-1
RED
YEL
BRN
ORG
WHI
BLK
J1
LINE
INPUT
10n
GND
LIFT
PREFERRED POSITION
R1
13k0
C1
620p
R$_L$*
A1
LM4562
OUTPUT
R2
4k99
R3
1k54

*R$_L$ MUST BE 10 kΩ OR GREATER
OMIT R1 & C1 (DAMPING NETWORK) FOR R$_L$ = 10 kΩ ONLY

LOW NOISE UNITY GAIN INPUT STAGE

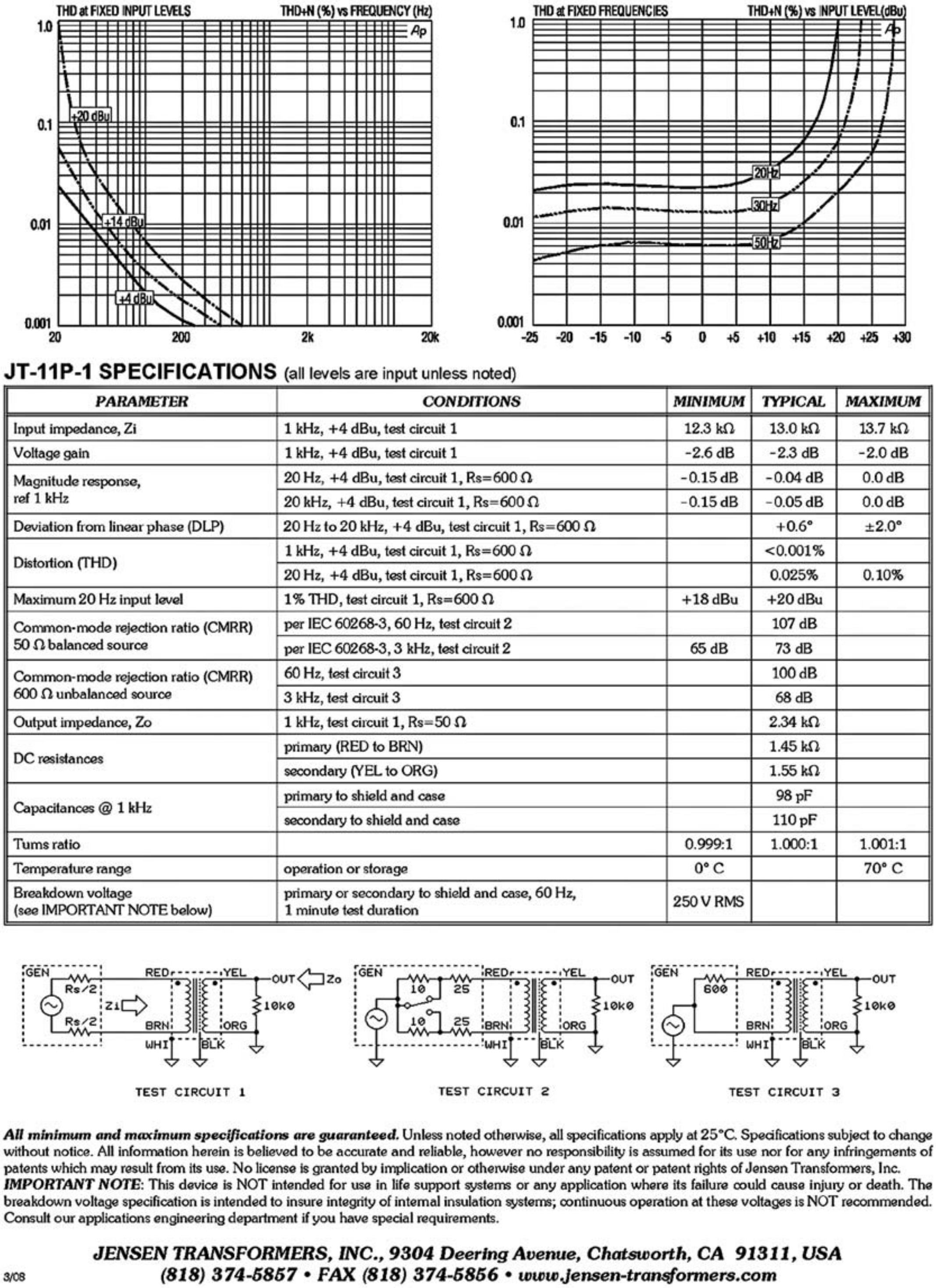
0.725 TYP
0.725 TYP
M
M
0.312 DIA TYP 2 PLACES
BOTTOM VIEW

USE ONLY #4 TYPE B SELF TAPPING SCREWS IN HOLES "M". ALLOW NO MORE THAN 0.15" PENETRATION INTO TRANSFORMER HOUSING.

TYPICAL APPLICATION

REFERENCE LEVEL RESPONSE
AMPLITUDE(dB) vs FREQUENCY(Hz)
Ap
+1
REF
-1
-2
-3
-4
-5
-6
0.2
2
20
200
2k
20k
200k
Low frequency data augmented with use of Hewlett-Packard 3325A Frequency Synthesizer
DEVIATION FROM LINEAR PHASE
DEVIATION (Deg) vs FREQUENCY (Hz)
Ap
+5
+4
+3
+2
+1
0
-1
-2
-3
-4
-5
20
200
2k
20k

Figure 11-55. Specification sheet for a quality transformer. Courtesy Jensen Transformers, Inc.


THD at FIXED INPUT LEVELS
THD+N (%) vs FREQUENCY (Hz)
Ap
1.0
0.1
0.01
0.001
20
200
2k
20k
+20 dBu
+14 dBu
+4 dBu
THD at FIXED FREQUENCIES
THD+N (%) vs INPUT LEVEL(dBu)
Ap
1.0
0.1
0.01
0.001
-25
-20
-15
-10
-5
0
+5
+10
+15
+20
+25
+30
20Hz
30Hz
50Hz

JT-11P-1 SPECIFICATIONS (all levels are input unless noted)

PARAMETER	*CONDITIONS*	*MINIMUM*	*TYPICAL*	*MAXIMUM*
Input impedance, Zi	1 kHz, +4 dBu, test circuit 1	12.3 kΩ	13.0 kΩ	13.7 kΩ
Voltage gain	1 kHz, +4 dBu, test circuit 1	-2.6 dB	-2.3 dB	-2.0 dB
Magnitude response, ref 1 kHz	20 Hz, +4 dBu, test circuit 1, Rs=600 Ω	-0.15 dB	-0.04 dB	0.0 dB
	20 kHz, +4 dBu, test circuit 1, Rs=600 Ω	-0.15 dB	-0.05 dB	0.0 dB
Deviation from linear phase (DLP)	20 Hz to 20 kHz, +4 dBu, test circuit 1, Rs=600 Ω		+0.6°	±2.0°
Distortion (THD)	1 kHz, +4 dBu, test circuit 1, Rs=600 Ω		<0.001%	
	20 Hz, +4 dBu, test circuit 1, Rs=600 Ω		0.025%	0.10%
Maximum 20 Hz input level	1% THD, test circuit 1, Rs=600 Ω	+18 dBu	+20 dBu	
Common-mode rejection ratio (CMRR) 50 Ω balanced source	per IEC 60268-3, 60 Hz, test circuit 2		107 dB	
	per IEC 60268-3, 3 kHz, test circuit 2	65 dB	73 dB	
Common-mode rejection ratio (CMRR) 600 Ω unbalanced source	60 Hz, test circuit 3		100 dB	
	3 kHz, test circuit 3		68 dB	
Output impedance, Zo	1 kHz, test circuit 1, Rs=50 Ω		2.34 kΩ	
DC resistances	primary (RED to BRN)		1.45 kΩ	
	secondary (YEL to ORG)		1.55 kΩ	
Capacitances @ 1 kHz	primary to shield and case		98 pF	
	secondary to shield and case		110 pF	
Turns ratio		0.999:1	1.000:1	1.001:1
Temperature range	operation or storage	0° C		70° C
Breakdown voltage (see IMPORTANT NOTE below)	primary or secondary to shield and case, 60 Hz, 1 minute test duration	250 V RMS		

GEN
Rs/2
Rs/2
Zi
RED
YEL
OUT
Zo
10kΩ
BRN
ORG
WHI
BLK
TEST CIRCUIT 1
GEN
10
25
10
25
RED
YEL
OUT
10kΩ
BRN
ORG
WHI
BLK
TEST CIRCUIT 2
GEN
600
RED
YEL
OUT
10kΩ
BRN
ORG
WHI
BLK
TEST CIRCUIT 3

All minimum and maximum specifications are guaranteed. Unless noted otherwise, all specifications apply at 25°C. Specifications subject to change without notice. All information herein is believed to be accurate and reliable, however no responsibility is assumed for its use nor for any infringements of patents which may result from its use. No license is granted by implication or otherwise under any patent or patent rights of Jensen Transformers, Inc. ***IMPORTANT NOTE***: This device is NOT intended for use in life support systems or any application where its failure could cause injury or death. The breakdown voltage specification is intended to insure integrity of internal insulation systems; continuous operation at these voltages is NOT recommended. Consult our applications engineering department if you have special requirements.

JENSEN TRANSFORMERS, INC., 9304 Deering Avenue, Chatsworth, CA 91311, USA
(818) 374-5857 • FAX (818) 374-5856 • www.jensen-transformers.com

3/08

Figure 11-55 Continued. Specification sheet for a quality transformer.

11.4 Installation and Maintenance

11.4.1 A Few Installation Tips

- Remember that there are very tiny wires inside an audio transformer. Its wire leads should never be used like a handle to pick it up. The internal bonds are strong, but pulling too hard might result in an open winding.
- Be careful with sharp tools. A gouge through the outer wrapper of an output transformer can nick or cut an internal winding.
- When mounting transformers that are in shielded cans, use either the supplied screws or ones no longer than recommended. If the screws are too long, they'll bore right into the windings—big problem!
- Be careful about using magnetized tools. If a screwdriver will pick up a paper clip, it shouldn't be used to install an audio transformer.
- Don't drop a transformer. It can distort the fit of the laminations in output transformers and affect their low-frequency response. Mechanical stress, as in denting of the magnetic shield can of an input transformer will reduce its effectiveness as a shield. For the same reason, don't over-tighten the clamps on transformers mounted with them.
- Twisting helps avoid hum pickup from ambient ac magnetic fields. This is especially true for microphone level lines in splitters, for example. Separately twist the leads from each winding—twisting the leads from *all* windings together can reduce noise rejection or CMRR.

11.4.2 De-Magnetization

Some subtle problems are created when transformer cores and/or their shield cans become magnetized. Generally, cores become magnetized by having dc flow in a winding, even for a fraction of a second. It can leave the core weakly magnetized. Steel cores, because of their wider hysteresis loops, are generally the most prone to such magnetization. The only way to know if the core has some permanent magnetization is to perform distortion measurements. A transformer with an un-magnetized core will exhibit nearly pure third harmonic distortion, with virtually no even order harmonic distortion while magnetized ones will show significant even order distortion, possibly with 2nd harmonic even exceeding 3rd. A test signal at a level about 30 or 40 dB below rated maximum operating level at 20 or 30 Hz is typically the most revealing because it maximizes the contribution of hysteresis distortion.

Microphone input transformers used with phantom power are exposed to this possibility whenever a microphone is connected or disconnected from a powered input. However, distortion tests before and after exposure to the worst-case 7 mA current pulses have shown that the effects are indeed subtle. Third harmonic distortion, which normally dominates transformer distortions, is unaffected. Second harmonic, which normally is near the measurement threshold, is typically increased by about 20 dB but is still some 15 dB lower than the third harmonic. Is it audible? Some say yes. But even this distortion disappears into the noise floor above a few hundred Hz. In any case, it can be prevented by connecting and disconnecting microphones *only* when phantom power is off. And such magnetized transformers can be de-magnetized.

Demagnetizing of low level transformers can generally be done with any audio generator having a continuously variable output. It may take a booster of some sort to get enough level for output transformers (be sure there's no dc offset at its output!). The idea is to drive the transformer deeply into saturation, 5% THD or more, and then slowly bring the level down to zero. Saturation will, of course, be easiest at a very low frequency. How much level it takes will depend on the transformer. If you're lucky, the level required may not be hazardous to the surrounding electronics and the de-magnetizing can be accomplished without disconnecting the transformer. Start with the generator set to 20 Hz and its minimum output level, connect it to the transformer, then slowly—over a period of a few seconds—increase the level into saturation—maintain it for a few seconds—then slowly turn it back down to minimum. For the vast majority of transformers, this process will leave them in a demagnetized state.

Shield cans are usually magnetized by having a brief encounter with a strongly magnetized tool. Sometimes, transformers are unknowingly mounted on a magnetized chassis. When the shield can of an input transformer becomes magnetized, the result is *microphonic* behavior of the transformer. Even though quality input transformers are potted with a semi-rigid epoxy compound to prevent breakage of very fine wires, vibration between core and can activate what is essentially a variable reluctance microphone. In this case, a good strong tape head de-magnetizer can be used to de-magnetize the can. At the end of the Jensen production line, most transformers are routinely demagnetized with a very strong de-magnetizer just prior to shipment. Although I haven't tried it, I would expect that something like a degausser for 2 inch video tape (remember that!) would also de-magnetize even a large steel-core output transformer.

References

1. Magnetic Shield Corporation, *Frequently Asked Questions*, www.magnetic-shield.com.
2. Sowter, G.A.V., *Soft Magnetic Materials for Audio Transformers: History, Production, and Applications*, Journal of the Audio Engineering Society, October 1987, www.sowter.co.uk/pdf/GAVS.pdf.
3. Whitlock, Bill, *Balanced Lines in Audio: Fact, Fiction, and Transformers*, Journal of the Audio Engineering Society, June 1995, pp 454-464.
4. Smith, F. Langford, *Radiotron Designer's Handbook*, Wireless Press, Sydney, 4th Edition, 1953, p 208.
5. Woolf, Lawrence, *RMS Watt, or Not?*, Electronics World, December 1998, pp 1043-1045.
6. Smith, F. Langford, *op. cit.*, p 227.

Notes

Co-Netic® is a registered trademark of Magnetic Shield Corp.

HyMu® is a registered trademark of Carpenter Technology Corp.

Mumetal® is a registered trademark of Telcon Metals, Ltd.

Permalloy® is a registered trademark of B & D Industrial & Mining Services, Inc.

This page intentionally left blank

Chapter 12

Tubes, Discrete Solid State Devices, and Integrated Circuits

by Glen Ballou

12.1 Tubes 311
  12.1.1 Tube Elements 311
  12.1.2 Tube Types 311
  12.1.3 Symbols and Base Diagrams 311
  12.1.4 Transconductance 312
  12.1.5 Amplification Factor 312
  12.1.6 Polarity 313
  12.1.7 Internal Capacitance 313
  12.1.8 Plate Resistance 314
  12.1.9 Grid Bias 314
  12.1.10 Plate Efficiency 315
  12.1.11 Power Sensitivity 315
  12.1.12 Screen Grid 315
  12.1.13 Plate Dissipation 315
  12.1.14 Changing Parameters 316
  12.1.15 Tube Heater 317
12.2 Discrete Solid-State Devices 317
  12.2.1 Semiconductors 317
  12.2.2 Diodes 320
  12.2.3 Thyristors 322
  12.2.4 Transistors 325
12.3 Integrated Circuits 334
  12.3.1 Monolithic Integrated Circuits 334
  12.3.2 Hybrid Integrated Circuits 335
  12.3.3 Operational Voltage Amplifiers (Op-Amp) 335
  12.3.4 Dedicated Analog Integrated Circuits for Audio Applications 340
    12.3.4.1 Voltage-Controlled Amplifiers 342
    12.3.4.2 Peak, Average, and RMS Level Detection 345
    12.3.4.3 Peak and Average Detection with Integrated Circuits 345
    12.3.4.4 Rms Level Detection Basics 345
    12.3.4.5 Rms Level Detection ICs 347
  12.3.5 Integrated Circuit Preamplifiers 349
    12.3.5.1 transformer input microphone preamplifiers 349
    12.3.5.2 Active Microphone Preamplifiers Eliminate Input Transformers 350
    12.3.5.3 The Evolution of Active Microphone Preamplifier ICs 350
    12.3.5.4 Integrated Circuit Mic Preamplifier Application Circuits 351

12.3.6 Balanced Line Interfaces. 353
  12.3.6.1 Balanced Line Inputs 354
  12.3.6.2 Balanced Line Outputs 354
  12.3.6.3 Integrated Circuits for Balanced Line Interfaces 355
  12.3.6.4 Balanced Line Input Application Circuits 356
  12.3.6.5 Balanced Line Receivers with the Common-Mode Performance of a Transformer 357
  12.3.6.6 InGenius High Common-Mode Rejection Line Receiver ICs 358
  12.3.6.7 Balanced Line Drivers 359
12.3.7 Digital Integrated Circuits. 360
References 364

12.1 Tubes

In 1883, Edison discovered that electrons flowed in an evacuated lamp bulb from a heated filament to a separate electrode (the Edison effect). Fleming, making use of this principle, invented the *Fleming valve* in 1905, but when DeForest, in 1907, inserted the grid, he opened the door to electronic amplification with the *audion*. The millions of vacuum tubes are an outgrowth of the principles set forth by these men.[1]

It was thought that, with the invention of the transistor and integrated circuits, the tube would disappear from audio circuits. This has hardly been the case. Recently tubes have had a revival because some "golden ears" like the smoothness and nature of the tube sound. The 1946 vintage 12AX7 is not dead and is still used today as are miniature tubes in condenser microphones and 6L6s in power amplifiers. It is interesting that many feel that a 50 W tube amplifier sounds better than a 250 W solid-state amplifier. For this reason, like the phonograph, tubes are still discussed in this handbook.

12.1.1 Tube Elements

Vacuum tubes consist of various elements or electrodes, Table 12-1. The symbols for these elements are shown in Fig. 12-1.

Table 12-1. Vacuum Tube Elements and Their Designation

Filament	The cathode in a directly heated tube that heats and emits electrons. A filament can also be a separate coiled element used to heat the cathode in an indirectly heated tube.
Cathode	The sleeve surrounding the heater that emits electrons. The surface of the cathode is coated with barium oxide or thoriated tungsten to increase the emission of electrons.
Plate	The positive element in a tube and the element from which the output signal is usually taken. It is also called an *anode*.
Control grid	The spiral wire element placed between the plate and cathode to which the input signal is generally applied. This element controls the flow of electrons or current between the cathode and the plate.
Screen grid	The element in a tetrode (four element) or pentode (five element) vacuum tube that is situated between the control grid and the plate. The screen grid is maintained at a positive potential to reduce the capacitance existing between the plate and the control grid. It acts as an electrostatic shield and prevents self-oscillation and feedback within the tube.
Suppressor grid	The gridlike element situated between the plate and screen in a tube to prevent secondary electrons emitted by the plate from striking the screen grid. The suppressor is generally connected to the ground or to the cathode circuit.

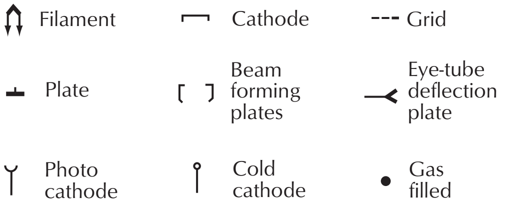


Figure 12-1. Tube elements and their designation.

12.1.2 Tube Types

There are many types of tubes, each used for a particular purpose. All tubes require a type of heater to permit the electrons to flow. Table 12-2 defines the various types of tubes.

Table 12-2. The Eight Types of Vacuum Tubes

Diode	A two-element tube consisting of a plate and a cathode. Diodes are used for rectifying or controlling the polarity of a signal as current can flow in one direction only.
Triode	A three-element tube consisting of a cathode, a control grid, and a plate. This is the simplest type of tube used to amplify a signal.
Tetrode	A four-element tube containing a cathode, a control grid, a screen grid, and a plate. It is frequently referred to as a screen-grid tube.
Pentode	A five-element tube containing a cathode, a control grid, a screen grid, a suppressor grid, and a plate.
Hexode	A six-element tube consisting of a cathode, a control grid, a suppressor grid, a screen grid, an injector grid, and a plate.
Heptode	A seven-element tube consisting of a cathode, a control grid, four other grids, and a plate.
Pentagrid	A seven-element tube consisting of a cathode, five grids, and a plate.
Beam-power tube	A power-output tube having the advantage of both the tetrode and pentode tubes. Beam-power tubes are capable of handling relatively high levels of output power for application in the output stage of an audio amplifier. The power-handling capabilities stem from the concentration of the plate-current electrons into beams of moving electrons. In the conventional tube the electrons flow from the cathode to the plate, but they are not confined to a beam. In a beam-power tube the internal elements consist of a cathode, a control grid, a screen grid, and two beam-forming elements that are tied internally to the cathode element. The cathode is indirectly heated as in the conventional tube.

12.1.3 Symbols and Base Diagrams

Table 12-3 gives the basic symbols used for tube circuits. The basing diagrams for various types of vacuum tubes are shown in Fig. 12-2.

Table 12-3. Tube Nomenclature

C	Coupling capacitor between stages
C_{g2}	Screen grid bypass capacitor
C_k	Cathode bypass capacitor
E_{bb}	Supply voltage
E_{ff}	Plate efficiency
E_p	Actual voltage at plate
E_{sg}	Actual voltage at screen grid
E_o	Output voltage
E_{sig}	Signal voltage at input
E_g	Voltage at control grid
E_f	Filament or heater voltage
I_f	Filament or heater current
I_p	Plate current
I_k	Cathode current
I_{sg}	Screen-grid current
I_{pa}	Average plate current
I_{pac}	Average ac plate current
I_{ka}	Average cathode current
I_{sga}	Average screen grid current
g_m	Transconductance (mutual conductance)
mu	Amplification factor (µ)
P_{sg}	Power at screen grid
P_p	Power at plate
P-P	Plate-to-plate or push–-pull amplifier
R_g	Grid resistor
R_k	Cathode resistor
R_l	Plate-load impedance or resistance
R_p	Plate-load resistor
R_{sg}	Screen-dropping resistor
R_d	Decoupling resistor
r_p	Internal plate resistance
V_g	Voltage gain

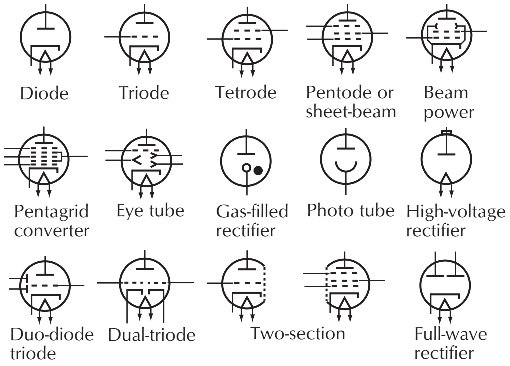


Figure 12-2. Basing diagrams for popular tubes.

12.1.4 Transconductance

Transconductance (g_m) is the change in the value of plate current expressed in microamperes (µA) divided by the signal voltage at the control grid of a tube, and is expressed by conductance. *Conductance* is the opposite of resistance, and the name *mho* (*ohm* spelled backward) was adopted for this unit of measurement. *Siemens* (S) have been adopted as the SI standard for conductance and are currently replacing mhos in measurement.

The basic mho or siemen is too large for practical usage; therefore, the terms *micromho* (µmho) and *microsiemens* (µS) are used. One micromho is equal to one-millionth of a mho.

The transconductance (g_m) of a tube in µmhos may be found with the equation

$$g_m = \frac{\Delta I_p}{\Delta E_{sig}} \tag{12-1}$$

where,
ΔI_p is the change of plate current,
ΔE_{sig} is the change of control-grid signal voltage,
E_{bb} is the plate supply voltage and is held constant.

For example, a change of 1 mA of plate current for a change of 1 V at the control grid is equal to a transconductance of 1000 µmho. A tube having a change of 2 mA plate current for a change of 1 V at the control grid would have a transconductance of 2000 µmho.

$$g_m = I_{pac} \times 1000 \tag{12-2}$$

where,
g_m is the transconductance in micromho or microsiemens,
I_{pac} is the ac plate current.

12.1.5 Amplification Factor

Amplification factor (µ) or *voltage gain* (V_g) is the ratio of the incremental plate voltage change to the control-electrode voltage change at a fixed plate current and constant voltage on all other electrodes. This normally is the amount the signal at the control grid is increased in amplitude after passing through the tube.

Tube voltage gain may be computed using the equation

$$V_g = \frac{\Delta E_p}{\Delta E_g} \tag{12-3}$$

where,
V_g is the voltage gain,

ΔE_p is the change in signal plate voltage,
ΔE_g is the change in the signal grid voltage.

If the amplifier consists of several stages, the amount of amplification is multiplied by each stage. The gain of an amplifier stage varies with the type of tube and the interstage coupling used. The general equation for voltage gain is

$$V_{gt} = V_{g1} V_{g2} \ldots V_{gn} \tag{12-4}$$

where,
V_{gt} is the total gain of the amplifier,
V_{g1}, V_{g2}, and V_{gn} are the voltage gain of the individual stages.

Triode tubes are classified by their amplification factor. A low-μ tube has an amplification factor less than 10. Medium-μ tubes have an amplification factor from 10–50, with a plate resistance of 5 Ω–15,000 Ω. High-μ tubes have an amplification factor of 50–100 with a plate resistance of 50 kΩ–100 kΩ.

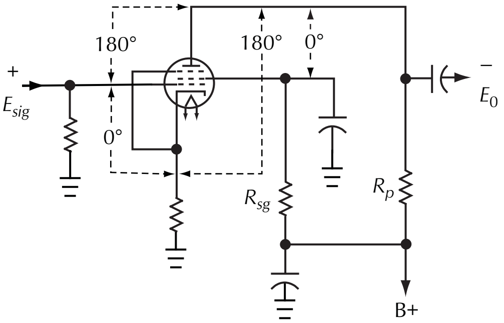
12.1.6 Polarity

Polarity reversals take place in a tube. The polarity reversal in electrical degrees between the elements of a self-biased pentode for a given signal at the control grid is shown in Fig. 12-3A. The reversals are the same for a triode. Note that, for an instantaneous positive voltage at the control grid, the voltage polarity between the grid and plate is 180° and will remain so for all normal operating conditions. The control grid and cathode are in polarity. The plate and screen-grid elements are in polarity with each other. The cathode is 180° out of polarity with the plate and screen-grid elements.

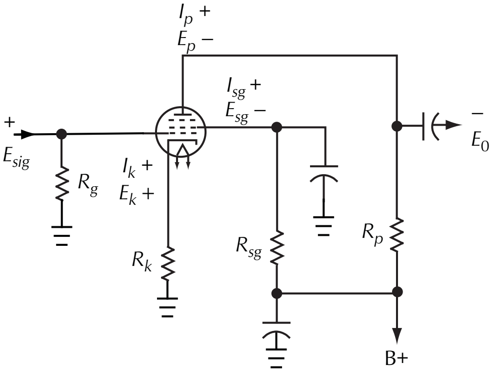
The polarity reversal of the instantaneous voltage and current for each element is shown in Fig. 12-3B. For an instantaneous positive sine wave at the control grid, the voltages at the plate and screen grid are negative, and the currents are positive. The voltage and current are both positive in the cathode resistor and are in polarity with the voltage at the control grid. The reversals are the same in a triode for a given element.

12.1.7 Internal Capacitance

The *internal capacitance* of a vacuum tube is created by the close proximity of the internal elements, Fig. 12-4. Unless otherwise stated by the manufacturer, the internal capacitance of a glass tube is measured using a close-fitting metal tube shield around the glass envelope connected to the cathode terminal. Generally, the capacitance is measured with the heater or filament cold and with no voltage applied to any of the other elements.



A. Polarity reversal of the signal between the elements of a pentode vacuum tube.



B. Polarity reversal of the current and voltage in a pentode vacuum tube.

Figure 12-3. Polarity characteristics of a vacuum tube.

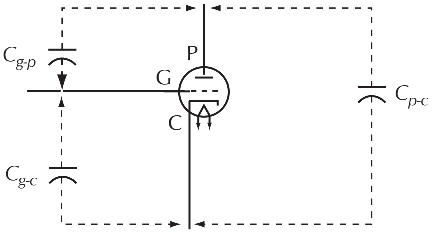


Figure 12-4. Interelectrode capacitance of a triode.

In measuring the capacitance, all metal parts, except the input and output elements, are connected to the cathode. These metal parts include internal and external shields, base sleeves, and unused pins. In testing a midsection tube, elements not common to the section being measured are connected to ground.

Input capacitance is measured from the control grid to all other elements, except the plate, which is connected to ground.

Output capacitance is measured from the plate to all other elements, except the control grid, which is connected to ground.

Grid-to-plate capacitance is measured from the control grid to the plate with all other elements connected to ground.

12.1.8 Plate Resistance

The *plate resistance* (r_p) of a vacuum tube is a constant and denotes the internal resistance of the tube or the opposition offered to the passage of electrons from the cathode to the plate. Plate resistance may be expressed in two ways: the dc resistance and the ac resistance. Dc resistance is the internal opposition to the current flow when steady values of voltage are applied to the tube elements and may be determined simply by using Ohm's Law

$$r_{p_{dc}} = \frac{E_p}{I_p} \tag{12-5}$$

where,

E_p is the dc plate voltage,

I_p is the steady value of plate current.

The ac resistance requires a family of plate-current curves from which the information may be extracted. As a rule, this information is included with the tube characteristics and is used when calculating or selecting components for an amplifier. The equation for calculating ac plate resistance is

$$r_{p_{ac}} = \frac{\Delta E_p}{\Delta I_p} \tag{12-6}$$

where,

ΔE_p is the change in voltage at the plate,

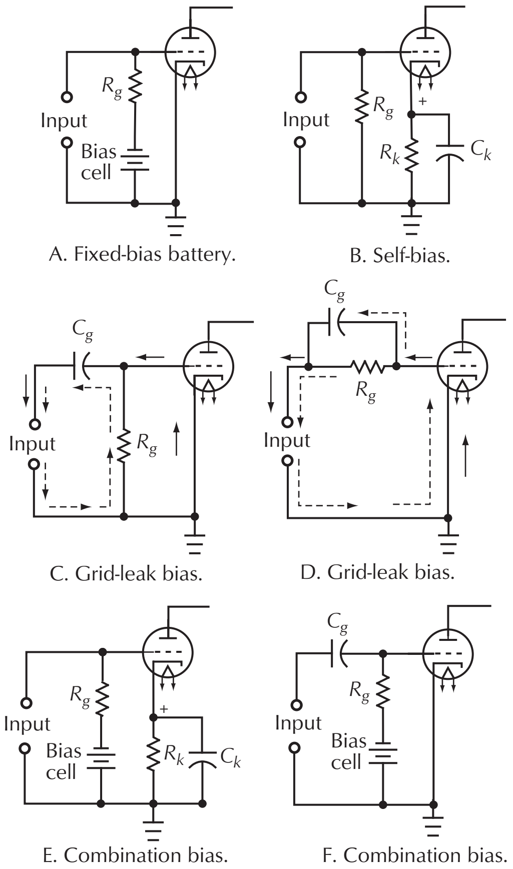
ΔI_p is the change in plate current,

E_{sig} is the control grid signal voltage and is held constant.

The values of E_p and I_p are taken from the family of curves supplied by the manufacturer for the particular tube under consideration.

12.1.9 Grid Bias

Increasing the plate voltage or decreasing the grid-bias voltage decreases the plate resistance. The six methods most commonly used to bias a tube are illustrated in Fig. 12-5. In Fig. 12-5A bias cell (battery) is connected in series with the control grid. In Fig. 12-5B the tube is self-biased by the use of a resistor connected in the cathode circuit. In Fig. 12-5C the circuit is also a form of self-bias; however, the bias voltage is obtained by the use of a grid capacitor and grid-leak resistor connected between the control grid and ground. In Fig. 12-5D the bias voltage is developed by a grid-leak resistor and capacitor in parallel, connected in series with the control grid. The method illustrated in Fig. 12-5E is called *combination bias* and consists of self-bias and battery bias. The resultant bias voltage is the negative voltage of the battery, and the bias created by the self-bias resistor in the cathode circuit. Another combination bias circuit is shown in Fig. 12-5F. The bias battery is connected in series with the grid-leak resistor. The bias voltage at the control grid is that developed by the battery and the self-bias created by the combination of the grid resistor and capacitor.



A. Fixed-bias battery. B. Self-bias.

C. Grid-leak bias. D. Grid-leak bias.

E. Combination bias. F. Combination bias.

Figure 12-5. Various methods of obtaining grid bias.

If the control grid becomes positive with respect to the cathode, it results in a flow of current between the control grid and the cathode through the external circuits. This condition is unavoidable because the wires of the control grid, having a positive charge, attract electrons passing from the cathode to the plate. It is important that the control-grid voltage is kept negative, reducing grid current and distortion.

Grid-current flow in a vacuum tube is generally thought of as being caused by driving the control grid into the positive region and causing the flow of grid current.

The grid voltage, plate-current characteristics are found through a series of curves supplied by the tube manufacturer, as shown in Fig. 12-6.

The curves indicate that for a given plate voltage the plate current and grid bias may be determined. For example, the manufacturer states that for a plate voltage of 250 V and a negative grid bias of –8 V, the plate current will be 9 mA, which is indicated at point A on the 250 V curve. If it is desired to operate this tube with a plate voltage of 150 V and still maintain a plate current of 9 mA, the grid bias will have to be changed to a –3 V.

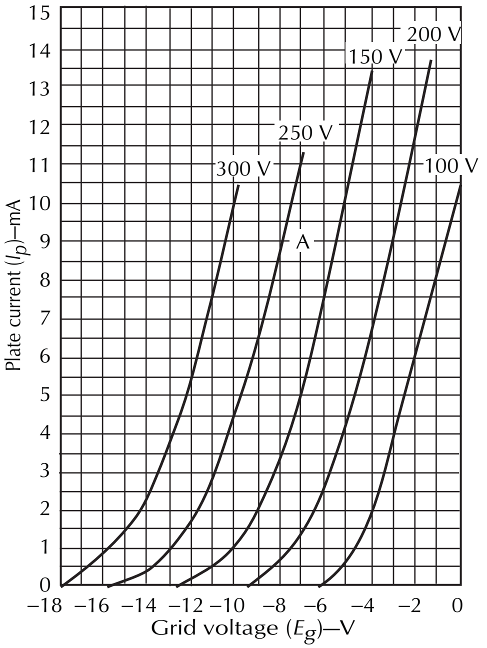


Figure 12-6. Grid voltage, plate-current curves for a triode tube.

12.1.10 Plate Efficiency

The *plate efficiency* (E_{ff}) is calculated by the equation:

$$E_{ff} = \frac{watts}{E_{pa}I_{pa}} \times 100 \qquad (12\text{-}7)$$

where,
watts is the power output,
E_{pa} is the average plate voltage,
I_{pa} is the average plate current.

The measurement is made with a load resistance in the plate circuit equal in value to the plate resistance stated by the manufacturer.

12.1.11 Power Sensitivity

Power sensitivity is the ratio of the power output to the square of the input voltage, expressed in mhos or siemens and is determined by the equation

$$\text{Power sensitivity} = \frac{P_o}{E_{in}^2} \qquad (12\text{-}8)$$

where,
P_o is the power output of the tube in watts,
E_{sig} is the rms signal voltage at the input.

12.1.12 Screen Grid

The *screen grid* series-dropping resistance is calculated by referring to the data sheet of the manufacturer and finding the maximum voltage that may be applied and the maximum power that may be dissipated by the screen grid. These limitations are generally shown graphically as in Fig. 12-7. The value of the resistor may be calculated using the equation

$$R_{sg} = \frac{E_{sg} \times (E_{bb} - E_{sg})}{P_{sg}} \qquad (12\text{-}9)$$

where,
R_{sg} is the minimum value for the screen-grid voltage-dropping resistor in ohms,
E_{sg} is the selected value of screen-grid voltage,
E_{bb} is the screen-grid supply voltage,
P_{sg} is the screen-grid input in watts corresponding to the selected value of E_{sg}.

12.1.13 Plate Dissipation

Plate dissipation is the maximum power that can be dissipated by the plate element before damage and is found with the equation

$$\text{Watts dissipation} = E_pI_p \qquad (12\text{-}10)$$

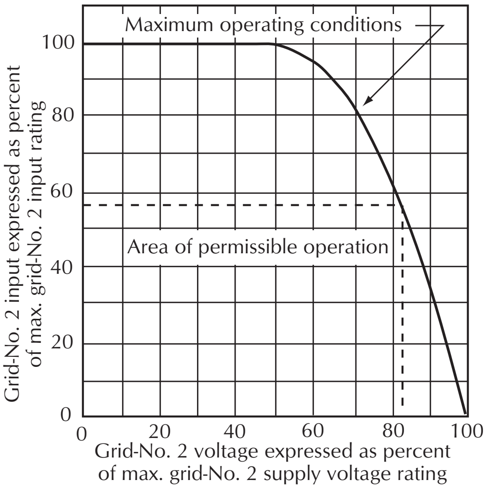


Figure 12-7. Typical graph for determining the maximum power dissipated by the screen grid.

where,
E_p is the voltage at the plate,
I_p is the plate current.

12.1.14 Changing Parameters

If a tube is to operate at a different plate voltage than published, the new values of bias, screen voltage, and plate resistance can be calculated by the use of conversion factors F_1, F_2, F_3, F_4, and F_5. Assume the following conditions are specified for a single beam-power tube:

Plate voltage	250.0 V
Screen voltage	250.0 V
Grid voltage	–12.5 V
Plate current	45.0 mA
Screen current	4.5 mA
Plate resistance	52,000.0 Ω
Plate load	5000.0 Ω
Transconductance	4100.0 µS
Power output	4.5 W

F_1 is used to find the new plate voltage

$$F_1 = \frac{E_{p_{new}}}{E_{p_{old}}} \tag{12-11}$$

For example, the new plate voltage is to be 180 V. The conversion factor F_1 for this voltage is obtained by dividing the new plate voltage by the published plate voltage Eq. 12-11:

$$F_1 = \frac{180}{250} = 0.72$$

The screen and grid voltage will be proportional to the plate voltage:

$$E_g = F_1 \times old\ grid\ voltage \tag{12-12}$$

$$E_{sg} = F_1 \times old\ screen\ voltage \tag{12-13}$$

In the example,

$$E_g = 0.72 \times (-12.5) = -9V$$

$$E_{sg} = 0.72 \times 250 = 180\text{ V}.$$

F_2 is used to calculate the plate and screen currents

$$F_2 = F_1\sqrt{F_1} \tag{12-14}$$

$$I_P = F_2 \times old\ plate\ current \tag{12-15}$$

$$I_s = F_2 \times old\ screen\ current \tag{12-16}$$

In the example,

$$F_2 = 0.72 \times 0.848 = 0.61$$

$$I_P = 0.61 \times 45\text{ mA} = 27.4\text{ mA}$$

$$I_{sg} = 0.61 \times 4.5\text{ mA} = 2.74\text{ mA}$$

The plate load and plate resistance may be calculated by use of factor F_3:

$$F_3 = \frac{F_1}{F_2} \tag{12-17}$$

$$r_p = F_3 \times old\ internal\ plate\ resistance \tag{12-18}$$

$$R_L = F_3 \times old\ \ plateload\ resistance \tag{12-19}$$

In the example,

$$F_3 = \frac{0.720}{0.610}$$
$$= 1.18$$

$$r_p = 1.18 \times 52{,}000$$
$$= 61{,}360\Omega$$

$$R_L = 1.18 \times 5000$$
$$= 5900\Omega$$

F_4 is used to find the power output

$$F_4 = F_1 F_2 \qquad (12\text{-}20)$$

$$Power\ output = F_4 \times old\ power\ output \qquad (12\text{-}21)$$

In the example:

$$F_4 = 0.72 \times 0.610$$
$$= 0.439$$

$$Power\ output = 0.439 \times 4.5$$
$$= 1.97\ \text{W}$$

F_5 is used to find the transconductance where

$$F_5 = \frac{1}{F_3} \qquad (12\text{-}22)$$

$$transconductance = F_5 \times old\ transconductance \qquad (12\text{-}23)$$

In the example,

$$F_5 = \frac{1}{1.18}$$
$$= 0.847$$

$$transconductance = 0.847 \times 4100$$
$$= (3472\ \mu\text{mho or } \mu\text{S})$$

The foregoing method of converting for voltages other than those originally specified may be used for triodes, tetrodes, pentodes, and beam-power tubes, provided the plate and grid 1 and grid 2 voltages are changed simultaneously by the same factor. This will apply to any class of tube operation, such as class A, AB_1, AB_2, B, or C. Although this method of conversion is quite satisfactory in most instances, the error will be increased as the conversion factor departs from unity. The most satisfactory region of operation will be between 0.7 and 2.0. When the factor falls outside this region, the accuracy of operation is reduced.

12.1.15 Tube Heater

The data sheets of tube manufacturers generally contain a warning that the heater voltage should be maintained within ±10% of the rated voltage. As a rule, this warning is taken lightly, and little attention is paid to heater voltage variations, which have a pronounced effect on the tube characteristics. Internal noise is the greatest offender. Because of heater-voltage variation, emission life is shortened, electrical leakage between elements is increased, heater-to-cathode leakage is increased, and grid current is caused to flow. Thus, the life of the tube is decreased with an increase of internal noise.

12.2 Discrete Solid-State Devices

12.2.1 Semiconductors

Conduction in solids was first observed by Munck and Henry in 1835 and later in 1874 by Braum. In 1905, Col. Dunwoody invented the crystal detector used in the detection of electromagnetic waves. It consisted of a bar of silicon carbide or carborundum held between two contacts. However, in 1903, Pickard filed a patent application for a crystal detector in which a fine wire was placed in contact with the silicon. This was the first mention of a silicon rectifier and was the forerunner of the present-day silicon rectifier. Later, other minerals such as galena (lead sulfide) were employed as detectors. During World War II, intensive research was conducted to improve crystal detectors used for microwave radar equipment. As a result of this research, the original point-contact transistor was invented at the Bell Telephone Laboratories in 1948.

A s*emiconductor* is an electronic device whose main functioning part is made from materials, such as germanium and silicon, whose conductivity ranges between that of a conductor and an insulator.

Germanium is a rare metal discovered by Winkler in Saxony, Germany, in 1896. Germanium is a by-product of zinc mining. Germanium crystals are grown from germanium dioxide powder. Germanium in its purest state behaves much like an insulator because it has very few electrical charge carriers. The conductivity of germanium may be increased by the addition of small amounts of an impurity.

Silicon is a nonmetallic element used in the manufacture of diode rectifiers and transistors. Its resistivity is considerably higher than that of germanium.

The relative position of pure germanium and silicon is given in Fig. 12-8. The scale indicates the resistance of conductors, semiconductors, and insulators per cubic centimeter. Pure germanium has a resistance of approximately 60 Ω/cm³. Germanium has a higher conductivity or less resistance to current flow than silicon and is used in low- and medium-power diodes and transistors.

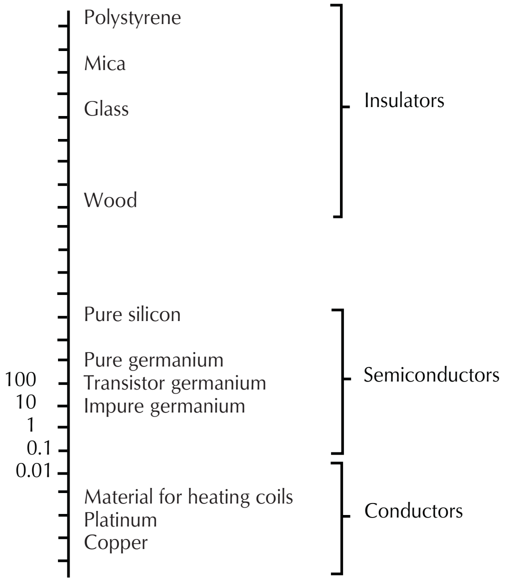


Figure 12-8. Resistance of various materials per cubic centimeter.

The base elements used to make semiconductor devices are not usable as semiconductors in their pure state. They must be subjected to a complex chemical, metallurgical, and photolithographical process wherein the base element is highly refined and then modified with the addition of specific impurities. This precisely controlled process of diffusing impurities into the pure base element is called *doping* and converts the pure base material into a semiconductor material. The semiconductor mechanism is achieved by the application of a voltage across the device with the proper polarity so as to have the device act either as an extremely low resistance (the forward biased or conducting mode) or as an extremely high resistance (reversed bias or nonconducting mode). Because the device is acting as both a good conductor of electricity and also, with the proper reversal of voltage, as a good electrical nonconductor or insulator, it is called a *semiconductor.*

Some semiconductor materials are called *p* or positive type because they are processed to have an excess of positively charged ions. Others are called *n* or negative type because they are processed to have an excess of negatively charged electrons. When a *p*-type of material is brought into contact with an n-type of material, a *pn junction* is formed. With the application of the proper external voltage, a low-resistance path is produced between the *n* and *p* material. By reversing the previously applied voltage, an extremely high-resistance called the *depletion layer* between the *p* and *n* types results. A diode is an example because its conduction depends upon the polarity of the externally applied voltage. Combining several of these *pn* junctions together in a single device produces semiconductors with extremely useful electrical properties.

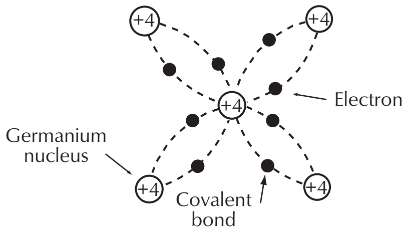
The theory of operation of a semiconductor device is approached from its atomic structure. The outer orbit of a germanium atom contains four electrons. The atomic structure for a pure germanium crystal is shown in Fig. 12-9A. Each atom containing four electrons forms covalent bonds with adjacent atoms, therefore, there are no "free" electrons. Germanium in its pure state is a poor conductor of electricity. If a piece of "pure" germanium (the size used in a transistor) has a voltage applied to it, only a few microamperes of current caused by electrons that have been broken away from their bonds by thermal agitation will flow in the circuit. This current will increase at an exponential rate with an increase of temperature.

When an atom with five electrons, such as antimony or arsenic, is introduced into the germanium crystal, the atomic structure is changed to that of Fig. 12-9B. The extra electrons (called free electrons) will move toward the positive terminal of the external voltage source.

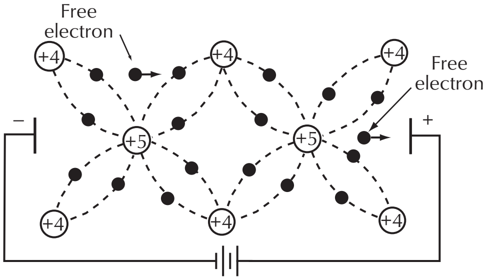
When an electron flows from the germanium crystal to the positive terminal of the external voltage source, another electron enters the crystal from the negative terminal of the voltage source. Thus, a continuous stream of electrons will flow as long as the external potential is maintained.

The atom containing the five electrons is the *doping agent* or *donor.* Such germanium crystals are classified as *n-type germanium.*

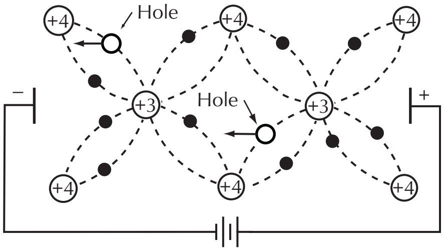
Using a doping agent of indium, gallium, or aluminum, each of which contains only three electrons in its outer orbit, causes the germanium crystal to take the atomic structure of Fig. 12-9C. In this structure, there is a *hole* or *acceptor.* The term *hole* is used to



A. Atomic structure of a pure germanium crystal. In this condition germanium is a poor conductor.



B. Atomic structure of an n-type germanium crystal when a doping agent containing five electrons is induced.



C. Atomic structure of a p-type germanium crystal when a doping agent containing three electrons is induced.

Figure 12-9. Atomic structure of germanium.

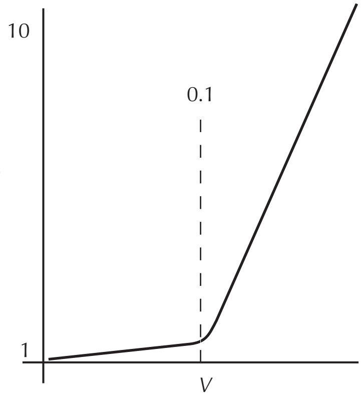
denote a mobile particle that has a positive charge and that simulates the properties of an electron having a positive charge.

When a germanium crystal containing holes is subjected to an electrical field, electrons jump into the holes, and the holes appear to move toward the negative terminal of the external voltage source.

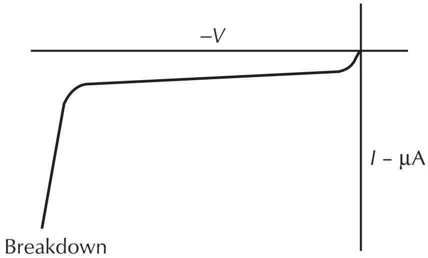
When a hole arrives at the negative terminal, an electron is emitted by the terminal, and the hole is canceled. Simultaneously, an electron from one of the covalent bonds flows into the positive terminal of the voltage source. This new hole moves toward the negative terminal causing a continuous flow of holes in the crystal.

Germanium crystals having a deficiency of electrons are classified *p-type germanium.* Insofar as the external electrical circuits are concerned, there is no difference between electron and hole current flow. However, the method of connection to the two types of transistors differs.

When a germanium crystal is doped so that it abruptly changes from an *n*-type to a *p*-type, and a positive potential is applied to the *p*-region, and a negative potential is applied to the *n*-region, the holes move through the junction to the right and the electrons move to the left, resulting in the voltage-current characteristic shown in Fig. 12-10A. If the potential is reversed, both electrons and holes move away from the junction until the electrical field produced by their displacement counteracts the applied electrical field. Under these conditions, zero current flows in the external circuit. Any minute amount of current that might flow is caused by thermal-generated hole pairs. Fig. 12-10B is a plot of the voltage versus current for the reversed condition. The leakage current is essentially independent of the applied potential up to the point where the junction breaks down.



A. Voltage-versus-current characteristic of the junction.



B. Voltage-versus-current characteristic of the junction transistor with the battery polarities in the reverse condition.

Figure 12-10. Voltage-versus-current characteristics.

12.2.2 Diodes

The *diode* is a device that exhibits a low resistance to current flow in one direction and a high resistance in the other. Ideally, when reverse biasing the diode (connecting the negative of the supply to the diode anode), no current should flow regardless of the value of voltage impressed across the diode. A forward-biased diode presents a very low resistance to current flow.

Fig. 12-11 shows the actual diode characteristics. Starting with the diode reverse biased, a small reverse current does flow. The size of this reverse-leakage current has been exaggerated for clarity and typically is in the order of nanoamperes. The forward resistance is not constant, and therefore it does not yield a straight-line forward-conduction curve. Instead, it begins high and drops rapidly at relatively low applied voltage. Above a 0.5–1 V drop it approaches a steep straight line slope (i.e., low resistance).

In the reverse-biased region of Fig. 12-11, when the applied voltage (–*V*) becomes large enough, the leakage current suddenly begins to increase very rapidly, and the slope of the characteristic curves becomes very steep. Past the *knee* in the characteristic, even a small increase in reverse voltage causes a large increase in the reverse current. This steep region is called the *breakdown* or *avalanche* region of the diode characteristic.

The application of high reverse voltage causes the diode to break down and stop behaving like a diode. *Peak-reverse-voltage* rating, or *prv* is one of the two most important diode parameters. This is also referred to as the *peak-inverse-voltage* rating, or *piv.* This rating indicates how high the reverse voltage can be without approaching the knee and risking breakdown. Additional diode parameters are:

Maximum average current	Causes overheating of the device
Peak repetitive current	Maximum peak value of current on a repetitive basis
Surge current	Absolute maximum allowed current even if just momentary

The maximum average current is limited by power dissipation in the junction. This power dissipation is represented by the product of forward voltage drop (V_F) and the forward current (I_F):

$$P = V_F I_F \tag{12-24}$$

Selenium Rectifiers and Diodes. A *selenium rectifier cell* consists of a nickel-plated aluminum baseplate coated with selenium, over which a low-temperature alloy is sprayed. The aluminum base serves as a negative electrode and the alloy as the positive. Current flows from the base plate to the alloy but encounters high resistance in the opposite direction. The efficiency of conversion depends to some extent on the ratio of the resistance in the conducting direction to that of the blocking direction. Conventional rectifiers generally have ratios from 100:1 to 1000:1.

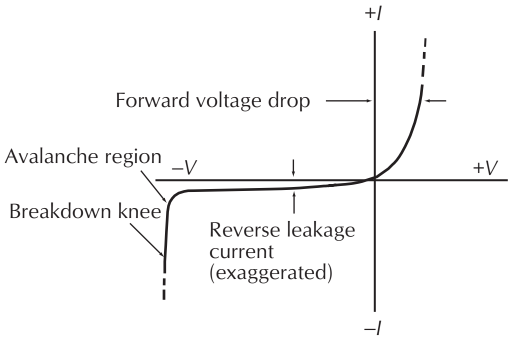


Figure 12-11. Actual diode characteristics.

Selenium rectifiers may be operated over temperatures of –55°C to +150°C (–67°F to +302°F). Rectification efficiency is on the order of 90% for three-phase bridge circuits and 70% for single-phase bridge circuits. As a selenium cell ages, the forward and reverse resistance increases for approximately one year and then stabilizes, decreasing the output voltage by approximately 15%. The internal impedance of a selenium rectifier is low and exhibits a nonlinear characteristic with respect to the applied voltage, maintaining a good voltage regulation. They are often used for battery charging.

Selenium rectifiers, because of their construction, have considerable internal capacitance which limits their operating range to audio frequencies. Approximate capacitance ranges are 0.10–0.15 µF/in^2 of rectifying surface.

The minimum voltage required for conduction in the forward direction is termed the *threshold voltage* and is about 1 V, therefore, selenium rectifiers cannot be used successfully below that voltage.

Silicon Rectifiers and Diodes. The high forward-to-reverse current characteristic of the silicon diode produces an efficiency of about 99%. When properly used, silicon diodes have long life and are not affected by aging, moisture, or temperature when used with the proper heat sink.

As an example, four individual diodes of 400 Vpiv may be connected in series to withstand a piv of 1600 V.

In a series arrangement, the most important consideration is that the applied voltage be equally distributed between the several units. The voltage drops across each individual unit must be very nearly identical. If the instantaneous voltage is not equally divided, one of the units may be subjected to a voltage exceeding its rated value, causing it to fail. This causes the other rectifiers to absorb the piv, often creating destruction of all the rectifiers.

Uniform voltage distribution can be obtained by the connection of capacitors or resistors in parallel with the individual rectifier unit, Fig. 12-12. Shunt resistors are used for steady-state applications, and shunt capacitors are used in applications where transient voltages are expected. If the circuit is exposed to both dc and ac, both shunt capacitors and resistors should be employed.

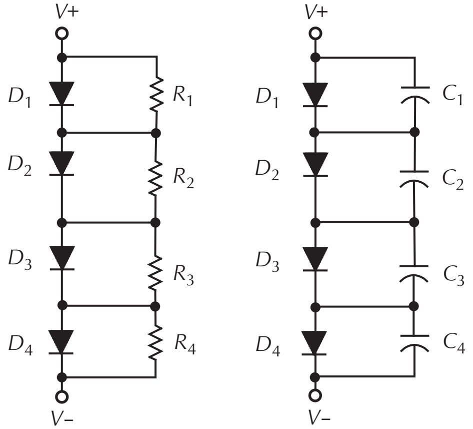


Figure 12-12. Rectifiers connected in series.

When the maximum current of a single diode is exceeded, two or more units may be connected in parallel. To avoid differences in voltage drop across the individual units, a resistor or small inductor is connected in series with each diode, Fig. 12-13. Of the two methods, the inductance is favored because of the lower voltage drop and consumption of power.

Zener and Avalanche Diodes. When the reverse voltage is increased beyond the breakdown knee of the diode characteristics as shown in Fig. 12-11, the diode impedance suddenly drops sharply to a very low value. If the current is limited by an external circuit resistance, operating in the "zener region" is normal for certain diodes specifically designed for the purpose. In *zener diodes*, sometimes simply called *zeners*, the breakdown characteristic is deliberately made as vertical as possible in the zener region so that the voltage across the diode is essentially constant over a wide reverse-current range, acting as a voltage regulator. Since its zener region voltage can be made highly repeatable and very stable with respect to time and temperature, the zener diode can also function as a voltage reference. Zener diodes come in a wide variety of voltages, currents, and powers, ranging from 3.2 V to hundreds of volts, from a few milliamperes to 10 A or more, and from about 250 mW to over 50 W.

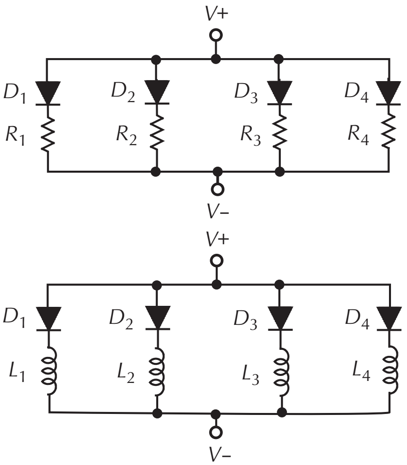


Figure 12-13. Rectifiers connected in parallel.

Avalanche diodes are diodes in which the shape of the breakdown knee has been controlled, and the leakage current before breakdown has been reduced so that the diode is especially well suited to two applications: high-voltage stacking and clamping. In other words, they prevent a circuit from exceeding a certain value of voltage by causing breakdown of the diode at or just below that voltage.

Small-Signal Diodes. *Small-signal diodes* or *general-purpose diodes* are low-level devices with the same general characteristics as power diodes. They are smaller, dissipate much less power, and are not designed for high-voltage, high-power operation. Typical rating ranges are:

I_F (forward current)	1–500 mA
V_F (forward voltage drop at I_F)	0.2–1.1 V
piv or prv	6–1000 V
I_R (leakage current at 80% prv)	0.1–1.0 μA

Switching Diodes. *Switching diodes* are small-signal diodes used primarily in digital-logic and control applications in which the voltages may change very rapidly so that speed, particularly reverse-recovery time, is of paramount importance. Other parameters of particular importance are low shunt capacitance, low and uniform V_F (forward voltage drop), low I_R (reverse leakage current), and in control circuits, prv.

Noise Diodes. *Noise diodes* are silicon diodes used in the avalanche mode (reverse biased beyond the breakdown knee) to generate broadband noise signals. All diodes generate some noise; these, however, have special internal geometry and are specially processed so as to generate uniform noise power over very broad bands. They are low-power devices (typically, 0.05–0.25 W) and are available in several different bandwidth classes from as low as 0 kHz–100 kHz to as high as 1000–18,000 MHz.

Varactor Diodes. *Varactor diodes* are made of silicon or gallium arsenide and are used as adjustable capacitors. Certain diodes, when operated in the reverse-biased mode at voltages below the breakdown value, exhibit a shunt capacitance that is inversely proportional to the applied voltage. By varying the applied reverse voltage, the capacitance of the varactor varies. This effect can be used to tune circuits, modulate oscillators, generate harmonics, and mix signals. Varactors are sometimes referred to as *voltage-tunable trimmer capacitors*.

Tunnel Diodes. The *tunnel diode* takes its name from the tunnel effect, a process where a particle can disappear from one side of a barrier and instantaneously reappear on the other side as though it had tunneled through the barrier element.

Tunnel diodes are made by heavily doping both the p and n materials with impurities, giving them a completely different voltage-current characteristic from regular diodes. This characteristic makes them uniquely useful in many high-frequency amplifiers as well as pulse generators and radiofrequency oscillators, Fig. 12-14.

What makes the tunnel diode work as an active element is the negative-resistance region over the voltage range V_d (a small fraction of a volt). In this region, increasing the voltage decreases the current, the opposite of what happens with a normal resistor. Tunnel diodes conduct heavily in the reverse direction; in fact, there is no breakdown knee or leakage region.

12.2.3 Thyristors

Stack four properly doped semiconductor layers in series, pnpn (or npnp), and the result is a four-layer, or Shockley breakover diode. Adding a terminal (gate) to the second layer creates a gate-controlled, reverse-blocking *thyristor*, or *silicon-controlled rectifier* (SCR), as shown in Fig. 12-15A.

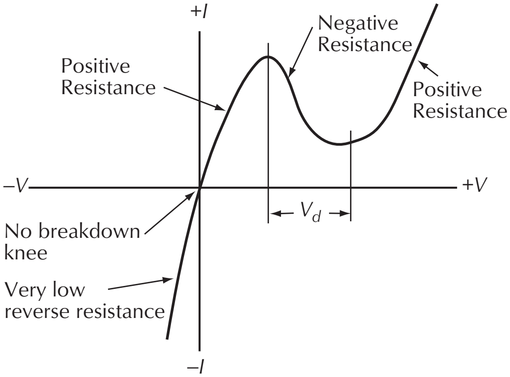
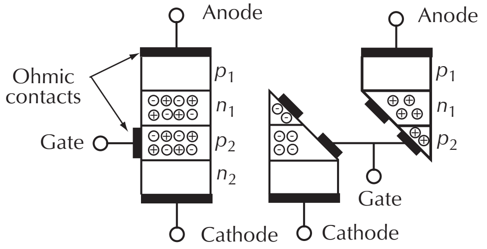
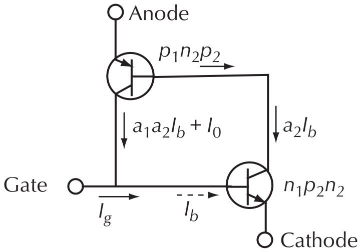


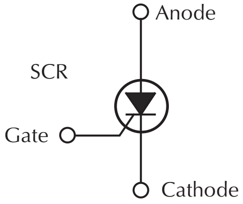
Figure 12-14. Tunnel-diode characteristics showing negative region (tunnel region).



A. Electrical layout of a thyristor.



B. Two-transistor equivalent circuit.



C. SCR layout.

Figure 12-15. Thyristor schematics.

The four-layer diode connects (fires) above a specific threshold voltage. In the SCR, the gate controls this firing threshold voltage, called the *forward blocking voltage*.

To understand how four-layer devices work, separate the material of the layers into two three-layer transistor devices. Fig. 12-15B is an equivalent two-transistor representation in a positive-feedback connection. Assuming a_1 and a_2 are the current gains of the two transistor sections with each gain value less than unity, the total base current I_b into the $n_1p_2n_2$ transistor is

$$I_b = a_1a_2I_b + I_o + I_g \tag{12-25}$$

where,

a_1 and a_2 are the transistor current gains,

I_b is the total base current,

I_o is the leakage current into the base of the $n_1p_2n_2$ transistor,

I_g is the current into the gate terminal.

The circuit turns on and becomes self-latching after a certain turn-on time needed to stabilize the feedback action, when the equality of Eq. 12-18 is achieved. This result becomes easier to understand by solving for I_b, which gives

$$I_b = \frac{I_o + I_g}{1 - a_1a_2} \tag{12-26}$$

When the product a_1a_2 is close to unity, the denominator approaches zero and I_b approaches a large value. For a given leakage current I_o, the gate current to fire the device can be extremely small. Moreover, as I_b becomes large, I_g can be removed, and the feedback will sustain the on condition since a_1 and a_2 then approach even closer to unity.

As applied anode voltage increases in the breakover diode, where I_g is absent, I_o also increases. When the quality of Eq. 12-18 is established, the diode fires. The thyristor fires when the gate current I_g rises to establish equality in the equation with the anode voltage fixed. For a fixed I_g, the anode voltage can be raised until the thyristor fires, with I_g determining the firing voltage, Fig. 12-16.

Once fired, a thyristor stays on until the anode current falls below a specified minimum holding current for a certain turnoff time. In addition, the gate loses all control once a thyristor fires. Removal or even reverse biasing of the gate signal will not turn off the device although reverse biasing can help speed turnoff. When the device is used with an ac voltage on the anode, the unit automatically turns off on the negative half of the voltage cycle. In dc switching circuits, however, complex means must often be used to remove, reduce, or reverse the anode voltage for turnoff.

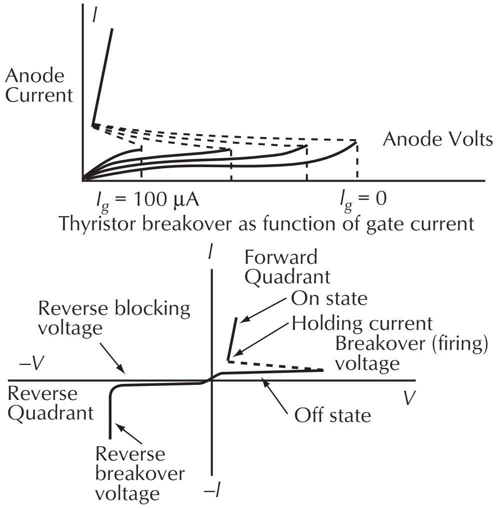


Figure 12-16. Thyristor breakover as a function of gate current and forward voltage.

Figure 12-17 shows a bilaterally conductive arrangement that behaves very much like two four-layer diodes (*diacs*), or two SCRs (*triacs*), parallel and oppositely conductive. When terminal A is positive and above the breakover voltage, a path through $p_1n_1p_2n_2$ can conduct; when terminal B is positive, path $p_2n_1p_1n_3$ can conduct. When terminal A is positive and a third element, terminal G, is sufficiently positive, the $p_1n_1p_2n_2$ path will fire at a much lower voltage than when G is zero. This action is almost identical with that of the SCR. When terminal G is made negative and terminal B is made positive, the firing point is lowered in the reverse, or $p_2n_1p_1n_3$, direction.

Because of low impedances in the on condition, four-layer devices must be operated with a series resistance in the anode and gate that is large enough to limit the anode-to-cathode or gate current to a safe value.

To understand the low-impedance, high-current capability of the thyristor, the device must be examined as a whole rather than by the two-transistor model. In Fig. 12-17B the $p_1n_1p_2$ transistor has holes injected to fire the unit, and the $n_1p_2n_2$ transistor has electrons injected. Considered separately as two transistors, the space-charge distributions would produce two typical transistor saturation-voltage forward drops, which are quite high when compared with the actual voltage drop of a thyristor.

However, when the thyristor shown in Fig. 12-17A is considered, the charges of both polarities exist simulta-

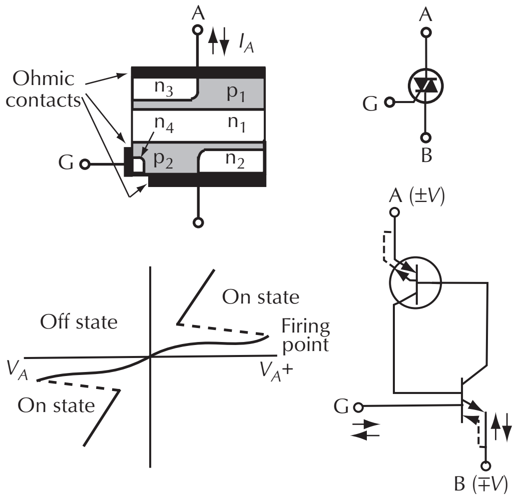


Figure 12-17. Bilateral arrangement to create a triac or ac operating device.

neously in the same n_1 and p_2 regions. Therefore, at the high injection levels that exist in thyristors, the mobile-carrier concentration of minority carriers far exceeds that from the background-doping density. Accordingly, the space charge is practically neutralized so that the forward drop becomes almost independent of the current density to high current levels. The major resistance to current comes from the ohmic contacts of the unit and load resistance.

The price paid for this low-impedance capability in a standard thyristor is a long turnoff time relative to turn-on time necessary to allow the high level of minority current carriers to dissipate. This long turnoff time limits the speed of a thyristor. Fortunately, this long turnoff time does not add significantly to switching power losses the way that a slow turnon time would.

Turnoff time is the minimum time between the forward anode current ceasing and the device being able to block reapplied forward voltage without turning on again.

Reverse-recovery time is the minimum time after forward conduction ceases that is needed to block reverse-voltage with ac applied to the anode-cathode circuit.

A third specification, *turnon time*, is the time a thyristor takes from the instant of triggering to when conduction is fully on.

These timing specifications limit the operating frequency of a thyristor. Two additional important specifications, the derivative of voltage with respect to time ($\mathrm{d}v/\mathrm{d}t$) and the derivative of current with respect to time ($\mathrm{d}i/\mathrm{d}t$) limit the rates of change of voltage and current application to thyristor terminals.

A rapidly varying anode voltage can cause a thyristor to turn on even though the voltage level never exceeds the forward breakdown voltage. Because of capacitance between the layers, a current large enough to cause firing can be generated in the gated layer. Current through a capacitor is directly proportional to the rate of change of the applied voltage; therefore, the $\mathrm{d}v/\mathrm{d}t$ of the anode voltage is an important thyristor specification.

Turnon by the $\mathrm{d}v/\mathrm{d}t$ can be accomplished with as little as a few volts per microsecond in some units, especially in older designs. Newer designs are often rated in tens to hundreds of volts per microsecond.

The other important rate effect is the anode-current $\mathrm{d}i/\mathrm{d}t$ rating. This rating is particularly important in circuits that have low inductance in the anode-cathode path. Adequate inductance would limit the rate of current rise when the device fires.

When a thyristor fires, the region near the gate conducts first; then the current spreads to the rest of the semiconductor material of the gate-controlled layer over a period of time. If the current flow through the device increases too rapidly during this period because the input-current $\mathrm{d}i/\mathrm{d}t$ is too high, the high concentration of current near the gate could damage the device do to localized overheating. Specially designed gate structures can speed up the turnon time of a thyristor, and thus its operational frequency, as well as alleviate this hot-spot problem.

Silicon-Controlled Rectifiers. The SCR thyristor can be considered a solid-state latching relay if dc is used as the supply voltage for the load. The gate current turns on the SCR, which is equivalent to closing the contacts in the load circuit.

If ac is used as the supply voltage, the SCR load current will reduce to zero as the positive ac wave shape crosses through zero and reverses its polarity to a negative voltage. This will shut off the SCR. If the positive gate voltage is also removed it will not turn on during the next positive half cycle of applied ac voltage unless positive gate voltage is applied.

The SCR is suitable for controlling large amounts of rectifier power by means of small gate currents. The ratio of the load current to the control current can be several thousand to one. For example, a 10 A load current might be triggered on by a 5 mA control current.

The major time-related specification associated with SCRs is the $\mathrm{d}v/\mathrm{d}t$ rating. This characteristic reveals how fast a transient spike on the power line can be before it false-triggers the SCR and starts its conducting without gate control current. Apart from this time-related parameter and its gate characteristics, SCR ratings are similar to those for power diodes.

SCRs can be used to control dc by using *commutating circuits* to shut them off. These are not needed on ac since the anode supply voltage reverses every half cycle. SCRs can be used in pairs or sets of pairs to generate ac from dc in inverters. They are also used as protective devices to protect against excessive voltage by acting as a short-circuit switch. These are commonly used in power supply *crowbar* overvoltage protection circuits. SCRs are also used to provide switched power-amplification, as in solid-state relays.

Triacs. The triac in Fig. 12-18 is a three-terminal semiconductor that behaves like two SCRs connected back to front in parallel so that they conduct power in both directions under control of a single gate-control circuit. Triacs are widely used to control ac power by phase shifting or delaying the gate-control signal for some fraction of the half cycle during which the power diode could be conducting. Light dimmers found in homes and offices and variable-speed drills are good examples of triac applications.

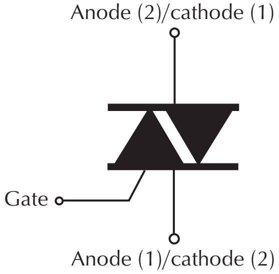


Figure 12-18. Schematic of a triac.

Light-Activated Silicon-Controlled Rectifiers. When sufficient light falls on the exposed gate junction, the SCR is turned on just as if the gate-control current were flowing. The gate terminal is also provided for optional use in some circuits. These devices are used in projector controls, positioning controls, photo relays, slave flashes, and security protection systems.

Diacs. The diac is shown in Fig. 12-19. It acts as two zener (or avalanche) diodes connected in series, back to back. When the voltage across the diac in either direction gets large enough, one of the zeners breaks down. The action drops the voltage to a lower level, causing a current increase in the associated circuit. This device is

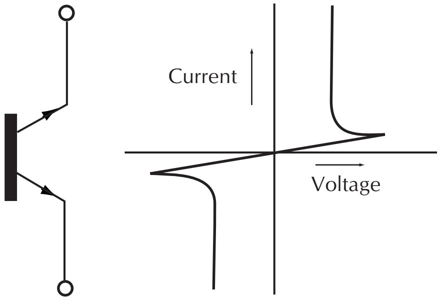


Figure 12-19. Schematic of a diac.

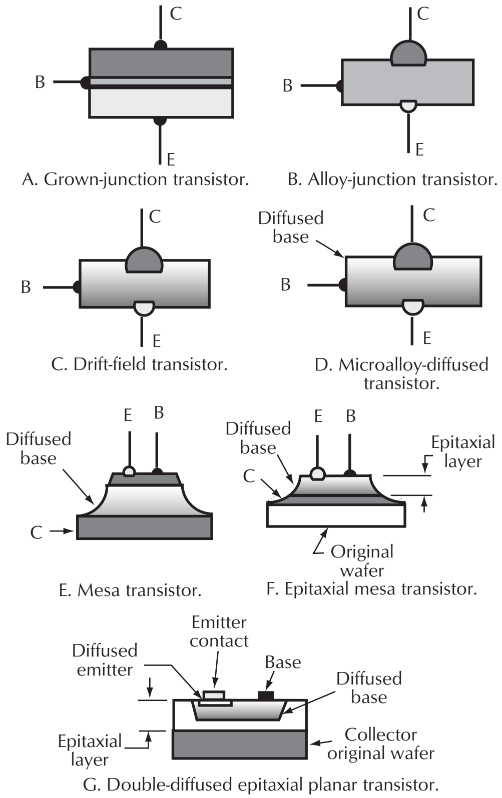
Opto-Coupled Silicon-Controlled Rectifiers. An *opto-coupled SCR* is a combination of a light-emitting diode (LED) and a photo silicon-controlled rectifier (photo-SCR). When sufficient current is forced through the LED, it emits an infrared radiation that triggers the gate of the photo-SCR. A small control current can regulate a large load current, and the device provides insulation and isolation between the control circuit (the LED) and the load circuit (the SCR). Opto-coupled transistors and Darlington transistors that operate on the same principle will be discussed later.

12.2.4 Transistors

There are many different types of transistors,[1] and they are named by the way they are grown, or made. Fig. 12-20A shows the construction of a *grown-junction* transistor. An *alloy-junction* transistor is shown in Fig. 12-20B. During the manufacture of the material for a grown junction, the impurity content of the semiconductor is altered to provide *npn* or *pnp* regions. The grown material is cut into small sections, and contacts are attached to the regions. In the alloy-junction type, small dots of *n*- or *p*-type impurity elements are attached to either side of a thin wafer of *p*- or *n*-type semiconductor material to form regions for the emitter and collector junctions. The base connection is made to the original semiconductor material.

Drift-field transistors, Fig. 12-20C, employ a modified alloy junction in which the impurity concentration in the wafer is diffused or graded. The drift field speeds up the current flow and extends the frequency response of the alloy-junction transistor. A variation of the drift-field transistor is the *microalloy diffused* transistor, as shown in Fig. 12-20D. Very narrow base dimensions are achieved by etching techniques, resulting in a shortened current path to the collector.

Mesa transistors shown in Fig. 12-20E use the original semiconductor material as the collector, with the



A. Grown-junction transistor. B. Alloy-junction transistor.

C. Drift-field transistor. D. Microalloy-diffused transistor.

E. Mesa transistor. F. Epitaxial mesa transistor.

G. Double-diffused epitaxial planar transistor.

Figure 12-20. Construction of various transistors.

base material diffused into the wafer and an emitter dot alloyed into the base region. A flat-topped peak or mesa is etched to reduce the area of the collector at the base junction. Mesa devices have large power-dissipation capabilities and can be operated at very high frequencies. *Double-diffused epitaxial mesa transistors* are grown by the use of vapor deposition to build up a crystal layer on a crystal wafer and will permit the precise control of the physical and electrical dimensions independently of the nature of the original wafer. This technique is shown in Fig. 12-20F.

The *planar transistor* is a highly sophisticated method of constructing transistors. A limited area source is used for both the base diffusion and emitter diffusion, which provides a very small active area, with a large wire contact area. The advantage of the planar construction is its high dissipation, lower leakage current, and lower collector cut-off current, which increases the stability and reliability. Planar construction is also used with several of the previously discussed base designs. A double-diffused epitaxial planar transistor is shown in Fig. 12-20G.

The *field-effect transistor*, or FET as it is commonly known, was developed by the Bell Telephone Laboratories in 1946, but it was not put to any practical use until about 1964. The principal difference between a conventional transistor and the FET is the transistor is a current-controlled device, while the FET is voltage controlled, similar to the vacuum tube. Conventional transistors also have a low-input impedance, which may at times complicate the circuit designer's problems. The FET has a high-input impedance with a low-output impedance, much like a vacuum tube.

The basic principles of the FET operation can best be explained by the simple mechanism of a *pn* junction. The control mechanism is the creation and control of a depletion layer, which is common to all reverse-biased junctions. Atoms in the *n* region possess excess electrons that are available for conduction, and the atoms in the *p* region have excess holes that may also allow current to flow. Reversing the voltage applied to the junction and allowing time for stabilization, very little current flows, but a rearrangement of the electrons and holes will occur. The positively charged holes will be drawn toward the negative terminals of the voltage source, and the electrons, which are negative, will be attracted to the positive terminal of the voltage source. This results in a region being formed near the center of the junction having a majority of the carriers removed and therefore called the *depletion regions*.

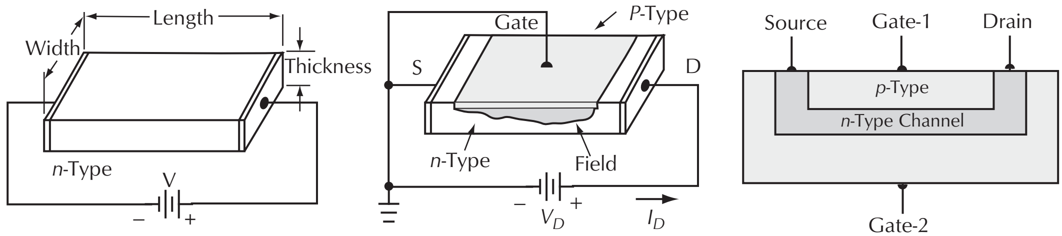
Referring to Fig. 12-21A, a simple bar composed of *n*-type semiconductor material has a nonrectifying contacts at each end. The resistance between the two end electrodes is

$$R = \frac{PL}{WT} \tag{12-27}$$

where,

P is the function of the material sensitivity,
L is the length of the bar,
W is the width,
T is the thickness.

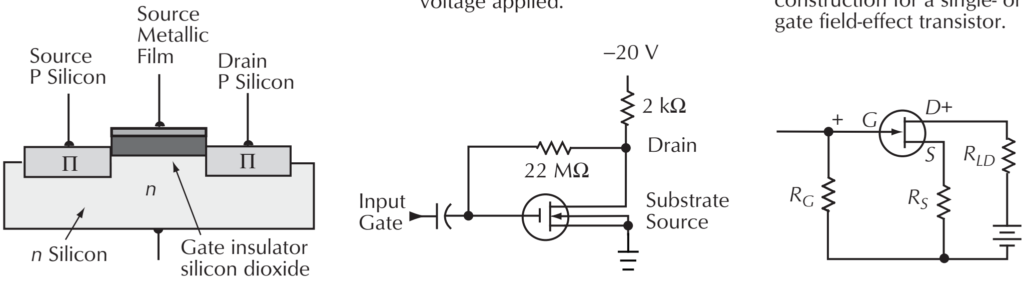
Varying one or more of the variables of the resistance of the semiconductor changes the bar. Assume a p-region in the form of a sheet is formed at the top of the bar shown in Fig. 12-21B. A *pn* junction is formed by diffusion, alloying, or epitaxial growth creating a reverse voltage between the *p* and *n*-material producing two depletion regions. Current in the n-material is caused primarily by means of excess electrons. By reducing the concentration of electrons or majority carriers, the resistivity of the material is increased. Removal of the excess electrons by means of the deple-



A. Plain semiconductor bar.

B. Bar with gate added and drain voltage applied.

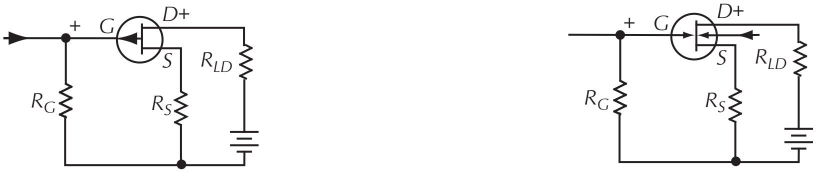
C. Cross-sectional view of the construction for a single- or double-gate field-effect transistor.



D. Internal construction of an insulated-gate transistor (IGT).

E. Typical circuit for an IGT transistor.

F. *n*-channel field-effect transistor circuit.



G. *p*-channel field-effect transistor circuit.

H. *n*-channel double-gate field-effect transistor circuit.

Figure 12-21. Field-effect transistors (FETs).

tion region causes the material to become practically nonconductive.

Disregarding the *p* region and applying a voltage to the ends of the bar cause a current and create a potential gradient along the length of the bar material, with the voltage increasing toward the right, with respect to the negative end or ground. Connecting the *p* region to ground causes varying amounts of reverse-bias voltage across the *pn* junction, with the greatest amount developed toward the right end of the *p* region. A reverse voltage across the bar will produce the same depletion regions. If the resistivity of the p-type material is made much smaller than that of the *n*-type material, the depletion region will then extend much farther into the *n* material than into the *p* material. To simplify the following explanation, the depletion of *p* material will be ignored.

The general shape of the depletion is that of a wedge, increasing the size from left to right. Since the resistivity of the bar material within the depletion area is increased, the effective thickness of the conducting portion of the bar becomes less and less, going from the end of the *p* region to the right end. The overall resistance of the semiconductor material is greater because the effective thickness is being reduced. Continuing to increase the voltage across the ends of the bar, a point is reached where the depletion region is extended practically all the way through the bar, reducing the effective thickness to zero. Increasing the voltage beyond this point produces little change in current.

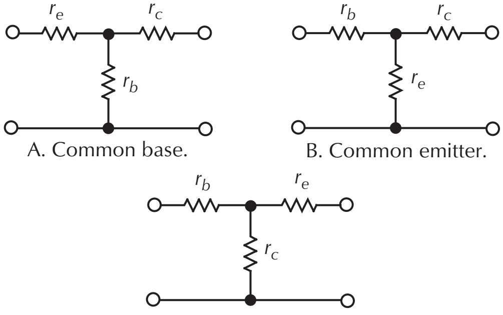
The *p* region controls the action and is termed a *gate*. The left end of the bar, being the source of majority carriers, is termed the *source*. The right end, being where the electrons are drained off, is called the *drain*. A cross-sectional drawing of a typical FET is shown in Fig. 12-21C, and three basic circuits are shown in Fig. 12-21F–H.

Insulated-gate transistors (IGT) are also known as field-effect transistors, metal-oxide silicon or semiconductor field-effect transistors (MOSFET), metal-oxide silicon or semiconductor transistors (MOST), and insulated-gate field-effect transistors (IGFET). All these devices are similar and are simply names applied to them by the different manufacturers.

The outstanding characteristics of the IGT are its extremely high input impedance, running to 10^{15} Ω. IGTs have three elements but four connections—the gate, the drain, the source, and an *n*-type substrate, into which two identical *p*-type silicon regions have been diffused. The source and drain terminals are taken from these two p regions, which form a capacitance between the *n* substrate and the silicon-dioxide insulator and the metallic gate terminals. A cross-sectional view of the internal construction appears in Fig. 12-21D, with a basic circuit shown in Fig. 12-21E. Because of the high input impedance, the IGT can easily be damaged by static charges. Strict adherence to the instructions of the manufacturer must be followed since the device can be damaged even before putting it into use.

IGTs are used in electrometers, logic circuits, and ultrasensitive electronic instruments. They should not be confused with the conventional FET used in audio equipment.

Transistor Equivalent Circuits, Current Flow, and Polarity. Transistors may be considered to be a T configuration active network, as shown in Fig. 12-22.



A. Common base.

B. Common emitter.

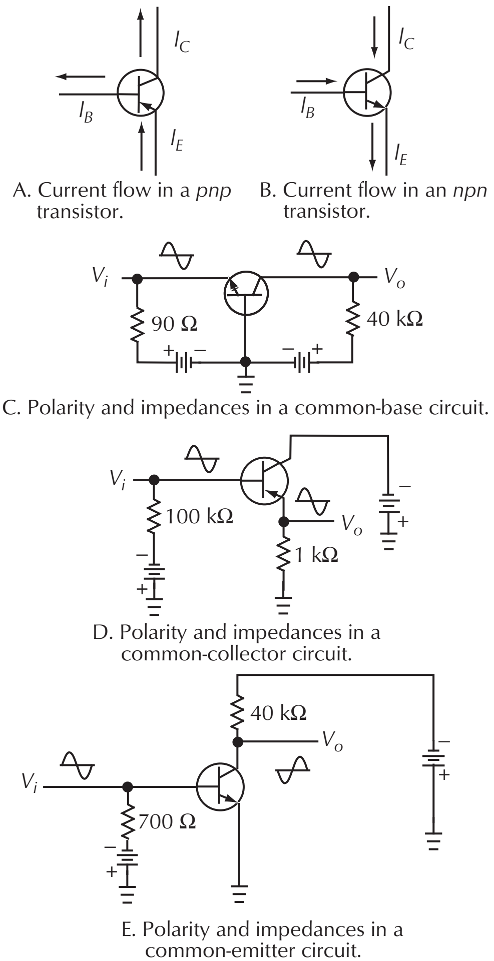
C. Common collector.

Figure 12-22. Equivalent circuits for transistors.

The current flow, phase, and impedances of the *npn* and *pnp* transistors are shown in Fig. 12-23 for the three basic configurations, common emitter, common base and common collector. Note phase reversal only takes place in the common-emitter configuration.

The input resistance for the common-collector and common-base configuration increases with an increase of the load resistance R_L. For the common emitter, the input resistance decreases as the load resistance is increased; therefore, changes of input or output resistance are reflected from one to the other.

Fig. 12-24 shows the signal-voltage polarities of a *p*-channel field-effect transistor. Note the similarity to tube characteristics.



A. Current flow in a *pnp* transistor.

B. Current flow in an *npn* transistor.

C. Polarity and impedances in a common-base circuit.

D. Polarity and impedances in a common-collector circuit.

E. Polarity and impedances in a common-emitter circuit.

Figure 12-23. Current, polarity. and impedance relationships.

Voltage, power, and current gains for a typical transistor using a common-emitter configuration are shown in Fig. 12-25. The current gain decreases as the load resistance is increased, and the voltage gain increases as the load resistance is increased. Maximum power gain occurs when the load resistance is approximately 40,000 Ω, and it may exceed unity.

For the common-collector connection, the current gain decreases as the load resistance is increased and the voltage gain increases as the load resistance is increased, but it never exceeds unity. Curves such as these help the designer to select a set of conditions for a specific result.

The power gain varies as the ratio of the input to output impedance and may be calculated with the equation

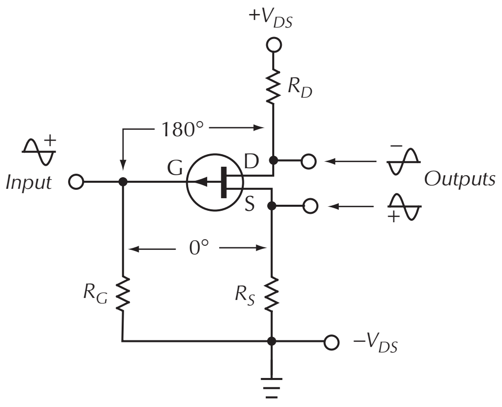


Figure 12-24. Signal-voltage polarities in a p-channel field-effect transistor (FET).

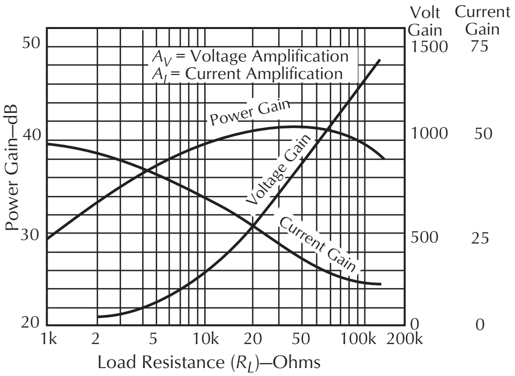


Figure 12-25. Typical voltage, power, and current gains for a conventional transistor using a common-emitter configuration.

$$dB = 10\log\frac{Z_o}{Z_{in}} \tag{12-28}$$

where,
Z_o is the output impedance in ohms,
Z_{in} is the input impedance in ohms.

Forward-Current-Transfer Ratio. An important characteristic of a transistor is its *forward-current-transfer ratio*, or the ratio of the current in the output to the current in the input element. Because of the many different configurations for connecting transistors, the forward transfer ratio is specified for a particular circuit configuration. The forward-current-transfer ratio for the common-base configuration is often referred to as *alpha* (α) and the common-emitter forward-current-transfer ratio as *beta* (β). In common-base circuitry, the emitter is the input element, and the collector is the output element. Therefore, α_{dc} is the ratio of the dc collector current I_C to the dc emitter current I_E. For the common emitter, the β_{dc} is then the ratio of the dc collector current I_C to the base current I_B. The ratios are also given in terms of the ratio of signal current, relative to the input and output, or in terms of ratio of change in the output current to the input current, which causes the change.

The terms α and β are also used to denote the frequency cutoff of a transistor and is defined as the frequency at which the value of α for a common-base configuration, or β for a common-emitter circuit, falls to 0.707 times its value at a frequency of 1000 Hz.

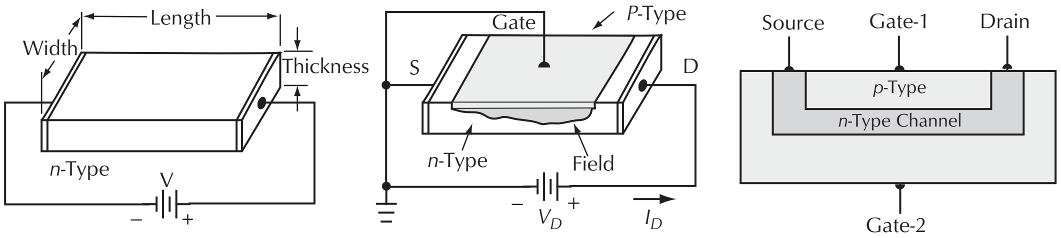
Gain-bandwidth product is the frequency at which the common-emitter forward-current-transfer ratio β is equal to unity. It indicates the useful frequency range of the device and assists in the determination of the most suitable configuration for a given application.

Bias Circuits. Several different methods of applying bias voltage to transistors are shown in Fig. 12-26, with a master circuit for aiding in the selection of the proper circuit shown in Fig. 12-27. Comparing the circuits shown in Fig. 12-26, their equivalents may be found by making the resistors in Fig. 12-27 equal to zero or infinity for analysis and study. As an example, the circuit of Fig. 12-26D may be duplicated in Fig. 12-27 by shorting out resistors R4 and R5 in Fig. 12-27.

The circuit Fig. 12-26G employs a split voltage divider for R_2. A capacitor connected at the junction of the two resistors shunts any ac feedback current to ground. The stability of circuits A, D, and G in Fig. 12-26 may be poor unless the voltage drop across the load resistor is at least one-third the value of the power supply voltage V_{cc}. The final determining factors will be gain and stability.

Stability may be enhanced by the use of a thermistor to compensate for increases in collector current with increasing temperature. The resistance of the thermistor decreases as the temperature increases, decreasing the bias voltage so the collector voltage tends to remain constant. Diode biasing may also be used for both temperature and voltage variations. The diode is used to establish the bias voltage, which sets the transistor idling current or the current flow in the quiescent state.

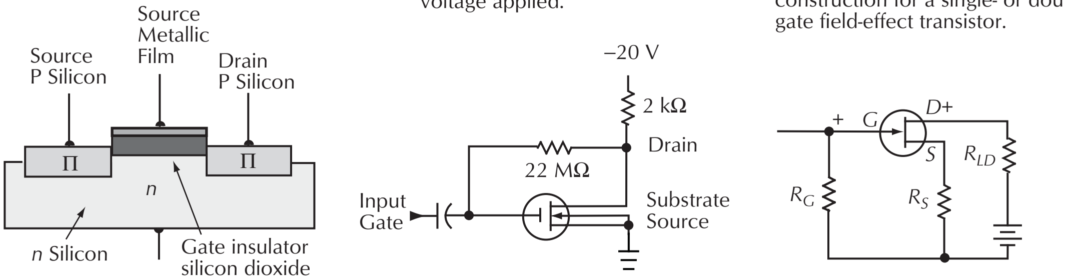
When a transistor is biased to a nonconducting state, small reverse dc currents flow, consisting of leakage currents that are related to the surface characteristics of the semiconductor material and saturation currents. Saturation current increases with temperature and is related to the impurity concentration in the material. Collector-cutoff current is a dc current caused when the collector-to-base circuit is reverse biased and the


Length
Width
Thickness
n-Type
V
–
+
Gate
P-Type
S
D
n-Type
Field
–
+
V_D
I_D
Source
Gate-1
Drain
p-Type
n-Type Channel
Gate-2

A. Plain semiconductor bar.

B. Bar with gate added and drain voltage applied.

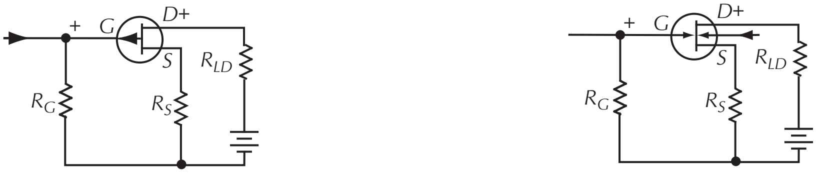
C. Cross-sectional view of the construction for a single- or double-gate field-effect transistor.


Source P Silicon
Source Metallic Film
Drain P Silicon
Π
Π
n
n Silicon
Gate insulator silicon dioxide
−20 V
2 kΩ
Drain
22 MΩ
Input Gate
Substrate
Source
+
G
D+
S
R_{LD}
R_G
R_S

D. Internal construction of an insulated-gate transistor (IGT).

E. Typical circuit for an IGT transistor.

F. *n*-channel field-effect transistor circuit.


+
G
D+
S
R_{LD}
R_G
R_S
+
G
D+
S
R_{LD}
R_G
R_S

G. *p*-channel field-effect transistor circuit.

H. *n*-channel double-gate field-effect transistor circuit.

Figure 12-26. Basic design circuit for transistor bias circuits.

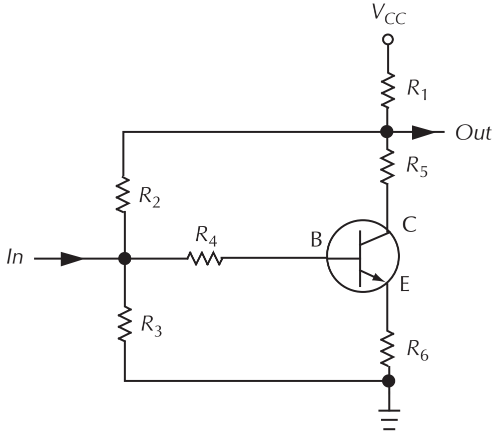

V_{CC}
R_1
Out
R_5
R_2
C
R_4
B
In
E
R_3
R_6

Figure 12-27. Basic bias circuits for transistors

emitter-to-base circuit is open. Emitter-cutoff current flows when the emitter to base is reverse biased and the collector-to-base circuit is open.

Small- and Large-Signal Characteristics. The transistor, like the vacuum tube, is nonlinear and can be classified as a nonlinear active device. Although the transistor is only slightly nonlinear, these nonlinearities become quite pronounced at very low and very high current and voltage levels. If an ac signal is applied to the base of a transistor without a bias voltage, conduction will take place on only one-half cycle of the applied signal voltage, resulting in a highly distorted output signal. To avoid high distortion, a dc-biased voltage is applied to the transistor, and the operating point is shifted to the linear portion of the characteristic curve. This improves the linearity and reduces the distortion to a value suitable for small-signal operation. Even though the transistor is biased to the most linear part of the characteristic curve, it can still add considerable distortion to the signal if driven into the nonlinear portion of the characteristic.

Small-signal swings generally run from less than 1 μV to about 10 mV so it is important that the dc-biased voltage be large enough that the applied ac signal is small compared to the dc bias current and voltage. Transistors are normally biased at current values between 0.1 mA and 10 mA. For large-signal operation, the design procedures become quite involved mathematically and require a considerable amount of approximation and the use of nonlinear circuit analysis.

It is important to provide an impedance match between cascaded stages because of the wide difference of impedance between the input and output circuits of transistors. If the impedances are not matched, an appreciable loss of power will take place.

The maximum power amplification is obtained with a transistor when the source impedance matches the internal input resistance, and the load impedance matches the internal output resistance. The transistor is then said to be image matched.

If the source impedance is changed, it affects the internal output resistance of the transistor, requiring a change in the value of the load impedance. When transistor stages are connected in tandem, except for the grounded-emitter connection, the input impedance is considerably lower than the preceding stage output impedance. Therefore, an interstage transformer should be used to supply an impedance match in both directions.

When working between a grounded base and a grounded-emitter circuit, a step-down transformer is used. Working into a grounded-collector stage, a step-up transformer is used. Grounded-collector stages can also be used as an impedance-matching device between other transistor stages.

When adjusting the supply voltages for a transistor amplifier employing transformers, the *battery* voltage must be increased to compensate for the dc voltage drop across the transformer windings. The data sheets of the manufacturer should be consulted before selecting a transformer to determine the source and load impedances.

Transistor Noise Figure (*nf*). In a low-level amplifier, such as a preamplifier, noise is the most important single factor and is stated as the SNR or *nf*. Most amplifiers employ resistors in the input circuit which contribute a certain amount of measurable noise because of thermal activity. This power is generally about –160 dB, re: 1 W, for a bandwidth of 10,000 Hz. When the input signal is amplified, the noise is also amplified. If the ratio of the signal power to noise power is the same, the amplifier is noiseless and has a noise figure of unity or more. In a practical amplifier some noise is present, and the degree of impairment is the noise figure (*nf*) of the amplifier, expressed as the ratio of signal power to noise power at the output:

$$nf = \frac{S_1 \times N_o}{S_o \times N_1} \tag{12-29}$$

where,
S_1 is the signal power,
N_1 is the noise power,
S_o is the signal power at the output,
N_o is the noise at the output.

$$nf_{dB} = 10\log(nf \text{ of the power ratio}) \tag{12-30}$$

For an amplifier with various *nf*, the SNR would be:

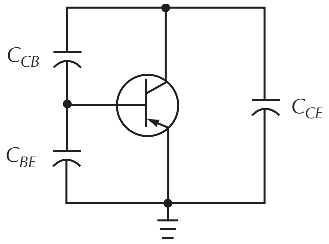
nf	SNR
1 dB	1.26
3 dB	2
10 dB	10
20 dB	100

An amplifier with an *nf* below 6 dB is considered excellent.

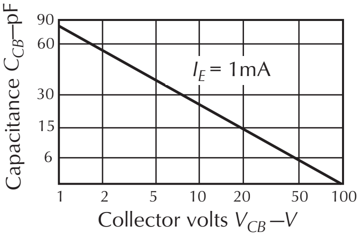
Low *nf* can be obtained by the use of an emitter current of less than 1 mA, a collector voltage of less than 2 V, and a signal-source resistance below 2000 Ω.

Internal Capacitance. The paths of internal capacitance in a typical transistor are shown in Fig. 12-28. The width of the *pn* junction in the transistor varies in accordance with voltage and current, and the internal capacitance also varies. Variation of collector-base capacitance *C* with collector voltage and emitter current is shown in Figs. 12-28B and C. The increase in the width of the *pn* junction between the base and collector, as the reverse bias voltage (V_{CB}) is increased, is reflected in lower capacitance values. This phenomenon is equivalent to increasing the spacing between the plates of a capacitor. An increase in the emitter current, most of which flows through the base-collector junction, increases the collector-base capacitance (C_{CB}). The increased current through the *pn* junction may be considered as effectively reducing the width of the pn junction. This is equivalent to decreasing the spacing between the plates of a capacitor, therefore increasing the capacitance.

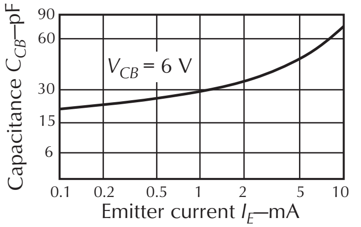
The average value of collector-base capacitance (C_{CB}) varies from 2–50 pF, depending on the type transistor and the manufacturing techniques. The collector-emitter capacitance is caused by the pn junction. It normally is five to ten times greater than that of the collector-base capacitance and will vary with the emitter current and collector voltage.



A. Capacitance between terminals.



B. Variation of C_{CB} with collector voltage.



C. Variation of C_{CB} with emitter current.

Figure 12-28. Internal capacitance of a transistor.

Punch-Through. *Punch-through* is the widening of the space charge between the collector element and the base of a transistor. As the potential V_{CB} is increased from a low to a high value, the collector-base space charge is widened. This widening effect of the space charge narrows the effective width of the base. If the diode space charge does not avalanche before the space charge spreads to the emitter section, a phenomenon termed punch-through is encountered, as shown in Fig. 12-29.

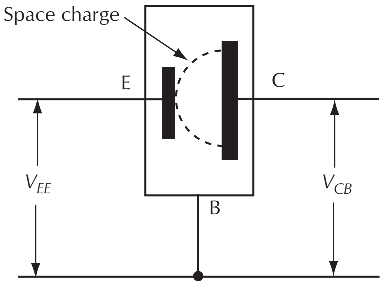


Figure 12-29. Spreading of the space charge between the emitter and the collector, which creates punch-through.

The effect is the base disappears as the collector-base space-charge layer contacts the emitter, creating relatively low resistance between the emitter and the collector. This causes a sharp rise in the current. The transistor action then ceases. Because there is no voltage breakdown in the transistor, it will start functioning again if the voltage is lowered to a value below where punch-through occurs.

When a transistor is operated in the punch-through region, its functioning is not normal, and heat is generated internally that can cause permanent damage to the transistor.

Breakdown Voltage. *Breakdown voltage* is that voltage value between two given elements in a transistor at which the crystal structure changes and current begins to increase rapidly. Breakdown voltage may be measured with the third electrode open, shorted, or biased in either the forward or reverse direction. A group of collector characteristics for different values of base bias are shown in Fig. 12-30. The collector-to-emitter breakdown voltage increases as the base-to-emitter bias is decreased from the normal forward values through zero to reverse. As the resistance in the base-to-emitter circuit decreases, the collector characteristics develop two breakdown points. After the initial breakdown, the collector-to-emitter voltage decreases with an increasing collector current, until another breakdown occurs at the lower voltage.

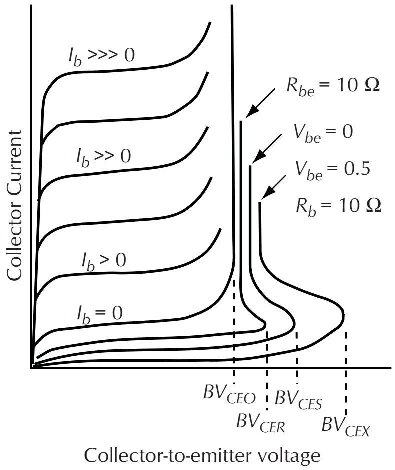


Figure 12-30. Typical collector characteristic curves showing locations of various breakdown voltages.

Breakdown can be very destructive in power transistors. A breakdown mechanism, termed *second breakdown*, is an electrical and thermal process in which current is concentrated in a very small area. The high current, together with the voltage across the transistor, causes intense heating, melting a hole from the collector to the emitter. This causes a short circuit and internal breakdown of the transistor.

The fundamental limitation to the use of transistors is the breakdown voltage (BV_{cer}). The breakdown voltage is not sharp so it is necessary to specify the value of collector current at which breakdown will occur. This data is obtained from the data sheet of the manufacturer.

Transistor Load Lines. Transistor load lines are used to design circuits. An example of circuit design uses a transistor with the following characteristics:

Maximum collector current	10 mA
Maximum collector voltage	–22 V
Base current	0 to 300 µA
Maximum power dissipation	300 mW

The base current curves are shown in Fig. 12-31A. The amplifier circuit is to be Class A, using a common-emitter circuit, as shown in Fig. 12-31B. By proper choice of the operating point, with respect to the transistor characteristics and supply voltage, low-distortion Class A performance is easily obtained within the transistor power ratings.

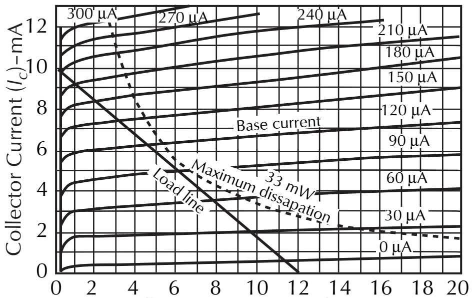
The first requirement is a set of collector-current, collector-voltage curves for the transistor to be employed. Such curves can generally be obtained from the data sheets of the manufacturer. Assuming that such data is at hand and referring to Fig. 12-31A, a curved line is plotted on the data sheet, representing the maximum power dissipation by the use of the equation

$$I_c = \frac{P_c}{V_c} \tag{12-31}$$

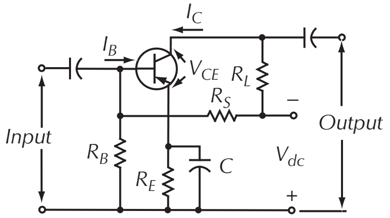
or

$$V_c = \frac{P_c}{I_c} \tag{12-32}$$

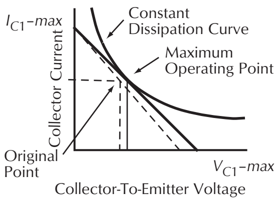
where,
I_c is the collector current,
P_c is the maximum power dissipation of the transistor,
V_c is the collector voltage.



A. Common-emitter-collector family of curves, with load line and maximum dissipation power curve.



B. Amplifier circuit used for load-line calculations.



C. Load line moved to right for maximum power output. Dotted lines are the original load line and operating point.

Figure 12-31. Load-line calculation curves.

At any point on this line at the intersection of V_cI_c, the product equals 0.033 W or 33 mW. In determining the points for the dissipation curve, voltages are selected along the horizontal axis and the corresponding current is equated using:

$$I_C = \frac{P_C}{V_{CE}} \tag{12-33}$$

The current is determined for each of the major collector-voltage points, starting at 16 V and working backward until the upper end of the power curve intersects the 300 µA base current line. After entering the value on the graph for the power dissipation curve, the area to the left of the curve encompasses all points within the maximum dissipation rating of the transistor. The area to the right of the curve is the overload region and is to be avoided.

The operating point is next determined. A point that results in less than a 33 mW dissipation is selected somewhere near the center of the power curve. For this example, a 5 mA collector current at 6 V, or a dissipation of 30 mW, will be used. The selected point is indicated on the graph and circled for reference. A line is drawn through the dot to the maximum collector current, 10 mA, and downward to intersect the V_{CE} line at the bottom of the graph, which, for this example, is 12 V. This line is termed the *load line*. The load resistance R_L may be computed with

$$R_L = \frac{dV_{CE}}{dI_C} \tag{12-34}$$

where,
R_L is the load resistance,
dV_{CE} is the range of collector-to-emitter voltage,
dI_C is the range of collector current.

In the example,

$$\begin{aligned} R_L &= \frac{0-12}{0-0.01} \\ &= \frac{12}{0.01} \\ &= 1200\Omega \end{aligned}$$

Under these conditions, the entire load line dissipates less than the maximum value of 33 mW, with 90 μA of base current and 5 mA of collector current. The required base current of 90 μA may be obtained by means of one of the biasing arrangements shown in Fig. 12-26.

To derive the maximum power output from the transistor, the load line may be moved to the right and the operating point placed in the maximum dissipation curve, as shown in Fig. 12-31C. Under these conditions, an increase in distortion may be expected. As the operating point is now at 6.5 V and 5 mA, the dissipation is 33 mW. Drawing a line through the new operating point and 10 mA (the maximum current), the voltage at the lower end of the load line is 13.0 V; therefore, the load impedance is now 1300 Ω.

12.3 Integrated Circuits

An *integrated circuit* (IC) is a device consisting of hundreds and even thousands of components in one small enclosure, and came into being when manufacturers learned how to grow and package semiconductors and resistors.

The first ICs were small scale and usually too noisy for audio circuits; however, as time passed, the noise was reduced, stability increased, and the operational amplifier (op-amp) IC became an important part of the audio circuit. With the introduction of medium-scale integration (MSI) and large-scale integration (LSI) circuits, power amplifiers were made on a single chip with only capacitors, gain, and frequency compensation components externally connected.

Typical circuit components might use up a space 4 mils × 6 mils (1 mil = 0.001 inch) for a transistor, 3 mils × 4 mils for a diode, and 2 mils × 12 mils for a resistor. These components are packed on the surface of the semiconductor wafer and interconnected by a metal pattern that is evaporated into the top surface. Leads are attached to the wafer that is then sealed and packaged in several configurations, depending on their complexity.

ICs can be categorized by their method of fabrication or use. The most common are monolithic or hybrid and linear or digital. Operational amplifiers and most analog circuits are linear while flip-flops and on–off switch circuits are digital.

An IC is considered *monolithic* if it is produced on one single chip and *hybrid* if it consists of more than one monolithic chip tied together and/or includes discrete components such as transistors, resistors, and capacitors.

With only a few external components, ICs can perform math functions, such as trigonometry, squaring, square roots, logarithms and antilogarithms, integration, and differentiation. ICs are well suited to act as voltage comparators, zero-crossing detectors, ac and dc amplifiers, audio and video amplifiers, null detectors, and sine-, square-, or triangular-wave generators, and all at a fraction of the cost of discrete-device circuits.

12.3.1 Monolithic Integrated Circuits

All circuit elements, both active and passive, are formed at the same time on a single wafer. The same circuit can be repeated many times on a single wafer and then cut to form individual 50 mil^2 ICs.

Bipolar transistors are often used in ICs and are fabricated much like the discrete transistor by the planar process. The differences are the contact-to-the-collector region is through the top surface rather than the substrate, requiring electrical isolation between the substrate and the collector. The integrated transistor is isolated from other components by a pn junction that creates capacitance, reducing high-frequency response and increasing leakage current, which in low-power circuits can be significant.

Integrated diodes are produced the same way as transistors and can be regarded as transistors whose terminals have been connected to give the desired characteristics.

Resistors are made at the same time as transistors. The resistance is characterized in terms of its sheet resistance, which is usually 100–200 Ω/square material for diffused resistors and 50–150 Ω/square material for deposited resistors. To increase the value of a resistor, square materials are simply connected in series.

It is very difficult to produce resistors with much closer tolerance than 10%; however, it is very easy to produce two adjacent resistors to be almost identical. When making comparator-type circuits, the circuits are balanced and are made to perform on ratios rather than absolute values. Another advantage is uniformity in temperature. As the temperature of one component varies, so does the temperature of the other components, allowing good tracking between components and circuits so ICs are usually more stable than discrete circuits.

Capacitors are made as thin-film integrated capacitors or junction capacitors. The thin-film integrated capacitor has a deposited metal layer and an *n*+ layer isolated with a carrier-free region of silicon dioxide. In junction capacitors, both layers are diffused low-resistance semiconductor materials. Each layer has a dopant of opposite polarity; therefore, the carrier-free region is formed by the charge-depleted area at the *pn* junction.

The MOSFET transistor has many advantages over the bipolar transistor for use in ICs as it occupies only $\frac{1}{25}$ the area of the bipolar equivalent due to lack of isolation pads. The MOSFET acts like a variable resistor and can be used as a high-value resistor. For instance, a 100 kΩ resistor might occupy only 1 mil^2 as opposed to 250 mil^2 for a diffused resistor.

The chip must finally be connected to terminals or have some means of connecting to other circuits, and it must also be packaged to protect it from the environment. Early methods included using fine gold wire to connect the chip to contacts. This was later replaced with aluminum wire ultrasonically bonded.

Flip-chip and beam-lead methods eliminate the problems of individually bonding wires. Relatively thick metal is deposited on the contact pads before the ICs are separated from the wafer. The deposited metal is then used to contact a matching metal pattern on the substrate. In the flip-chip method, globules of solder deposited on each contact pad ultrasonically bond the chip to the substrate.

In the beam-lead method, thin metal tabs lead away from the chip at each contact pad. The bonding of the leads to the substrate reduces heat transfer into the chip and eliminates pressure on the chip.

The chip is finally packaged in either hermetically sealed metal headers or is encapsulated in plastic, which is an inexpensive method of producing ICs.

12.3.2 Hybrid Integrated Circuits

Hybrid circuits combine monolithic and thick- and thin-film discrete components for obtaining the best solution to the design.

Active components are usually formed as monolithics; however, sometimes discrete transistors are soldered into the hybrid circuit.

Passive components such as resistors and capacitors are made with thin- and thick-film techniques. Thin films are 0.001–0.1 mil thick, while thick films are normally 60 mils thick. Resistors can be made with a value from ohms to megohms with a tolerance of 0.05% or better.

High-value capacitors are generally discrete, miniature components that are welded or soldered into the circuit, and low-value capacitors can be made as film capacitors and fabricated directly on the substrate.

Along with being certain that the components will fit into the hybrid package, the temperature must also be taken into account. The temperature rise T_R of the package can be calculated with the following equation:

$$\begin{aligned} T_R &= T_C - T_A \\ &= P_T T \theta_{CA} \end{aligned} \tag{12-35}$$

where,
T_C is the case temperature,
T_A is the ambient temperature,
P_T is the total power dissipation,
θ_{CA} is the case-to-ambient thermal resistance.

The θ_{CA} for a package in free air can be approximated at 35°C/W/in^2 or a device will have a 35°C rise in temperature above ambient if 1 W is dissipated over an area of 1 in^2.

12.3.3 Operational Voltage Amplifiers (Op-Amp)

One of the most useful ICs for audio is the op-amp. Op-amps can be made with discrete components, but they would be very large and normally unstable to temperature and external noise.

An op-amp normally has one or more of the following features:

- Very high input impedance ($>10^6$–10^{12} Ω).
- Very high open-loop (no feedback) gain.

- Low output impedance (<200 Ω).
- Wide frequency response (>100 MHz).
- Low input noise.
- High symmetrical slew rate and/or high input dynamic range.
- Low inherent distortion.

By adding external feedback paths, gain, frequency response, and stability can be controlled.

Op-amps are normally two-input differential devices; one input inverting the signal and the second input not inverting the signal, and hence called *noninverting*. Several typical op-amp circuits are shown in Fig. 12-32.

Because there are two inputs of opposite polarity, the output voltage is the difference between the inputs where

$$E_{O(+)} = A_V E_2 \tag{12-36}$$

$$E_{O(-)} = A_V E_1 \tag{12-37}$$

E_O is calculated with the equation

$$E_O = A_V \times (E_1 - E_2) \tag{12-38}$$

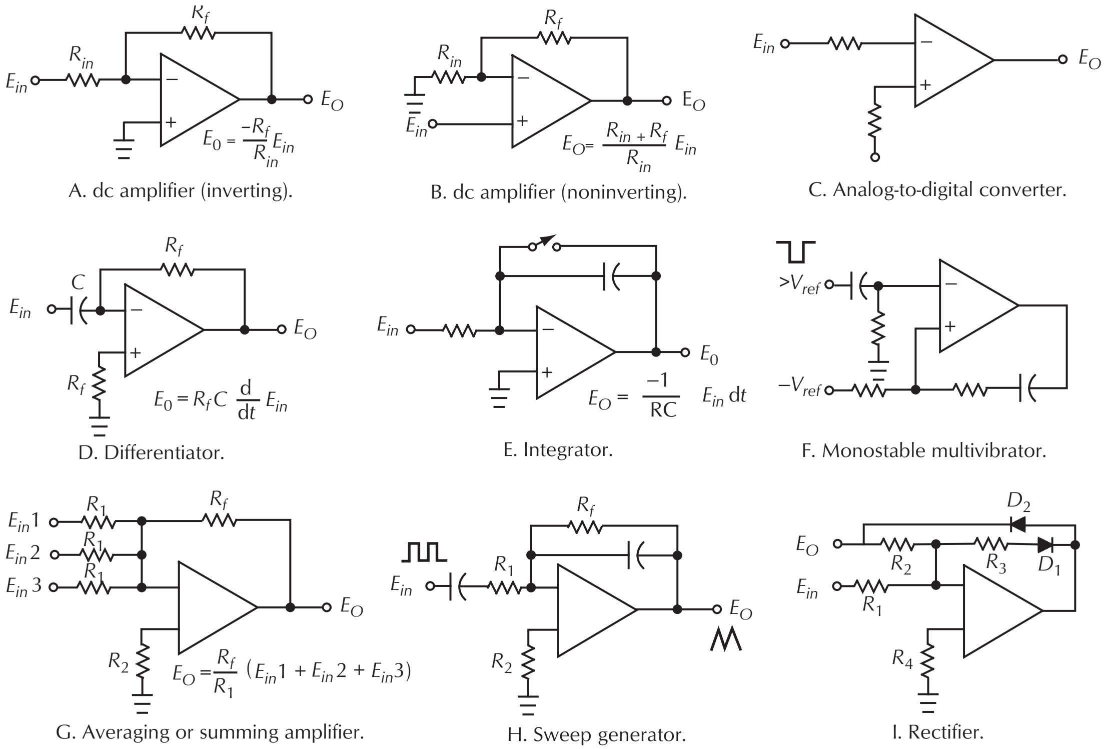
Often one of the inputs is grounded, either through a direct short or a capacitor. Therefore, the gain is either

$$E_O = A_V E_1 \tag{12-39}$$

or

$$E_O = A_V E_2 \tag{12-40}$$

To provide both a positive and negative output with respect to ground, a positive and negative power supply is required, as shown in Fig. 12-33. The supply should be regulated and filtered. Often a + and – power supply is not available, such as in an automobile, so the op-amp must operate on a single supply, as shown in Fig. 12-34. In this supply, the output dc voltage is set by adjusting R_1 and R_2 so the voltage at the noninverting input is about one-third the power supply voltage.



R_f
R_{in}
E_{in}
E_O
$E_0 = \frac{-R_f}{R_{in}} E_{in}$
$E_O = \frac{R_{in} + R_f}{R_{in}} E_{in}$
C
$E_0 = R_f C \frac{d}{dt} E_{in}$
$E_O = \frac{-1}{RC} E_{in}\,dt$
$>V_{ref}$
$-V_{ref}$
$E_{in}1$
$E_{in}2$
$E_{in}3$
R_1
R_2
R_3
R_4
D_1
D_2
$E_O = \frac{R_f}{R_1}(E_{in}1 + E_{in}2 + E_{in}3)$

A. dc amplifier (inverting).

B. dc amplifier (noninverting).

C. Analog-to-digital converter.

D. Differentiator.

E. Integrator.

F. Monostable multivibrator.

G. Averaging or summing amplifier.

H. Sweep generator.

I. Rectifier.

Figure 12-32. Typical op-amp circuits.

The diodes and zener diodes in Fig. 12-35 are used to protect the op-amp from damage caused by transients, reverse voltage, and overdriving. D_6 and D_7 clip the inputs before overdriving, D_1 and D_2 protect against reverse polarity, D_4 and D_5 regulate the supply, and D_3 limits the total voltage across the op-amp.

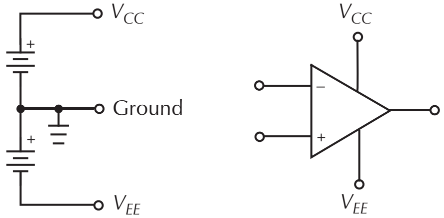

V_{CC}
V_{CC}
+
Ground
+
V_{EE}
V_{EE}

Figure 12-33. Positive- and negative-type power supply.

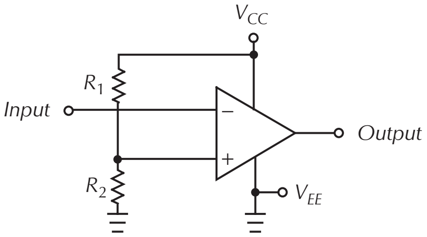

V_{CC}
R_1
Input
Output
R_2
V_{EE}

Figure 12-34. Simple circuit for operating on a single-ended power supply.

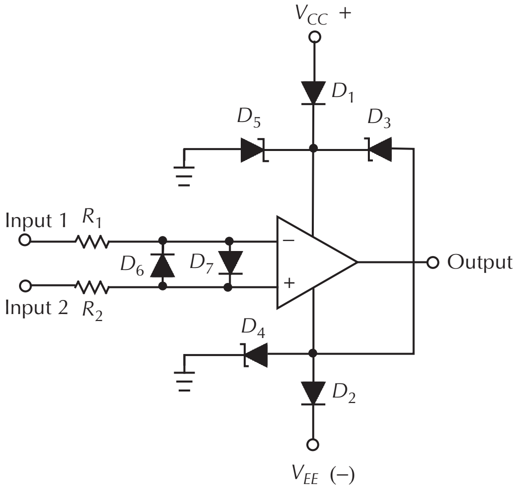

V_{CC} +
D_1
D_5
D_3
Input 1
R_1
D_6
D_7
Output
Input 2
R_2
D_4
D_2
V_{EE} (–)

Figure 12-35. Diode protection circuits for op-amps.

The dc error factors result in an output offset voltage E_{Oo}, which exists between the output and ground when it should be zero. The dc offset error is most easily corrected by supplying a voltage differential between the inverting and noninverting inputs, which can be accomplished by one of several methods, Fig. 12-36. Connecting the feedback resistor R_f usually causes an offset and can be found with the equation

$$E_{Oo} = I_{bias}R_f \tag{12-41}$$

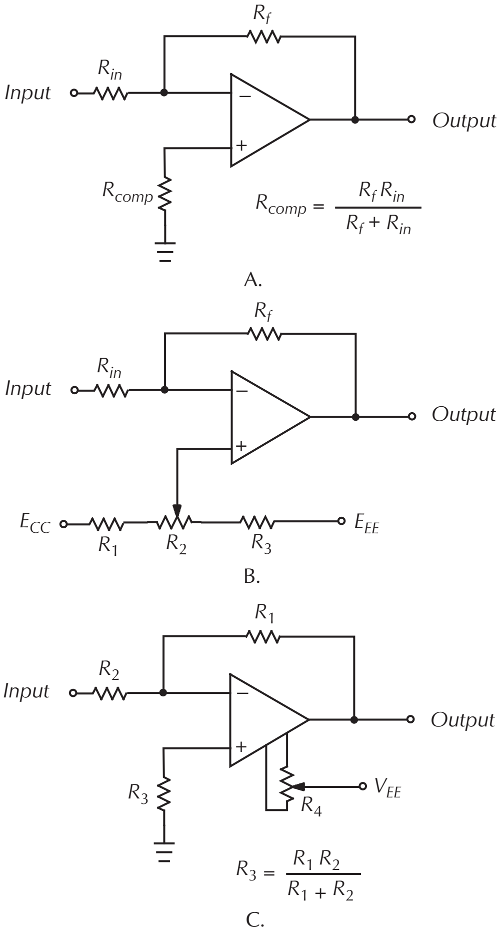

R_f
R_{in}
Input
Output
R_{comp}
$R_{comp} = \frac{R_f R_{in}}{R_f + R_{in}}$
A.
R_f
R_{in}
Input
Output
E_{CC}
R_1
R_2
R_3
E_{EE}
B.
R_1
R_2
Input
Output
R_3
R_4
V_{EE}
$R_3 = \frac{R_1 R_2}{R_1 + R_2}$
C.

Figure 12-36. Various methods of correcting dc error.

To obtain minimum offset, make the compensating resistor shown in Fig. 12-36A equal to

$$R_{comp} = \frac{R_f R_{in}}{R_f + R_{in}} \tag{12-42}$$

If this method is not satisfactory, the methods of Figs. 12-36B or C might be required.

Many op-amps are internally compensated. Often it is advantageous to compensate a device externally to optimize bandwidth and slew rate, lowering distortion. Internally compensated op-amp ICs come in standard packages—the 8 pin TO-99 metal can, the 8 pin dual-in-line package (MINI DIP), and the 14 pin DIP.

Inverting Amplifiers. In the *inverting amplifier* the + input is grounded and the signal is applied to the minus (–) input, Fig. 12-37. The output of the circuit is determined by the input resistor R_1 and the feedback resistor R_f.

$$E_O = E_{in}\left(\frac{-R_f}{R_1}\right) \tag{12-43}$$

where,
E_{in} is the signal input voltage in volts,
R_f is the feedback resistor in ohms,
R_1 is the input resistor in ohms.

The low frequency rolloff is

$$f_C = \frac{1}{2\pi R_1 C_1} \tag{12-44}$$

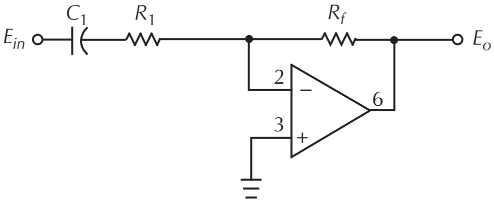


Figure 12-37. A simple inverting amplifier.

Noninverting Amplifier. In the *noninverting amplifier,* Fig. 12-38. the signal is applied to the plus input, while the minus input is part of the feedback loop. The output is

$$E_O = I_{in}\left(\frac{1+R_f}{R_1}\right) \tag{12-45}$$

The low-frequency rolloff is in two steps.

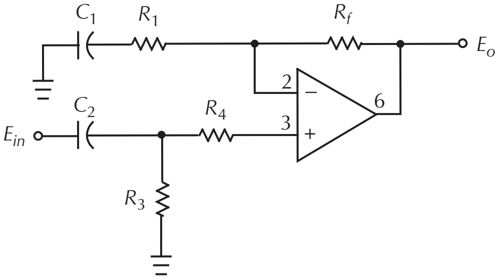


Figure 12-38. A simple noninverting amplifier.

$$f_{C_1} = \frac{1}{2\pi R_1 C_1} \tag{12-46}$$

$$f_{C_2} = \frac{1}{2\pi R_3 C_2} \tag{12-47}$$

To keep low-frequency noise gain at a minimum, keep $f_{C_1} > f_{C_2}$.

Power Supply Compensation. The power supply for wideband op-amp circuits should be bypassed with capacitors, Fig. 12-39A, between the plus and minus pin and common. The leads should be as short as possible and as close to the IC as possible. If this is not possible, bypass capacitors should be on each printed circuit board.

Input Capacitance Compensation. Stray input capacitance can lead to oscillation in feedback op-amps because it represents a potential phase shift at the frequency of

$$f = \frac{1}{2\pi R_f C_s} \tag{12-48}$$

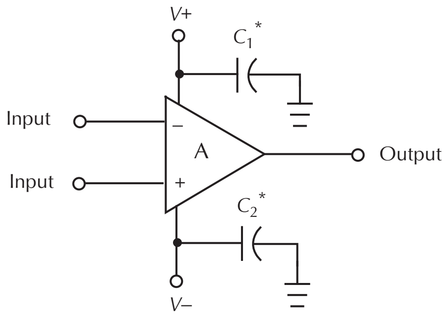
where,
R_f is the feedback resistor,
C_s is the stray capacitance.

One way to reduce this problem is to keep the value of R_f low. The most useful way, however, is to add a compensation capacitor, C_f, across R_f as shown in Fig. 12-39B. This makes C_f/R_f and C_s/R_{in} a frequency compensated divider.

Output Capacitance Compensation. Output capacitance greater than 100 pF can cause problems, requiring a series resistor R_o being installed between the output of the IC and the load and stray capacitance as shown in Fig. 12-39C. The feedback resistor (R_f) is connected after R_o to compensate for the loss in signal caused by R_o. A compensating capacitor (C_f) bypasses R_f to reduce gain at high frequencies.

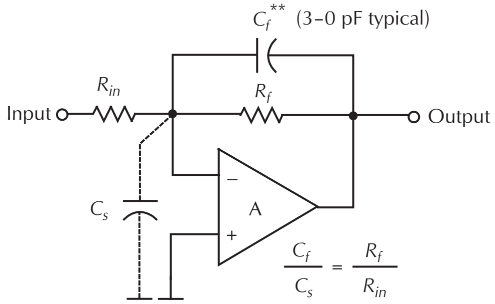
Gain and Bandwidth. A perfect op-amp would have infinite gain and infinite bandwidth. In real life however, the dc open loop voltage gain is around 100,000 or 100 dB and the bandwidth where gain is 0 is 1 MHz, Fig. 12-40.

To determine the gain possible in an op-amp, for a particular bandwidth, determine the bandwidth, follow vertically up to the open loop gain response curve and horizontally to the voltage gain. This, of course, is with no feedback at the upper frequency. For example, for a


V+
C1*
Input
Input
A
Output
C2*
V–

*Low-inductance short-lead capacitors—0.1 μF stacked film preferred. For high-speed op amps, connect C_1 and C_2 directly at supply pins, with low-inductance ground returns.

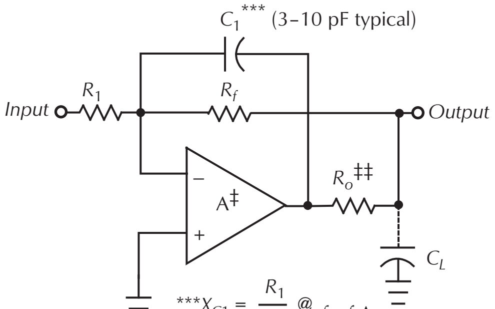
A. Power-supply bypassing.


Cf** (3–0 pF typical)
Rin
Rf
Input
Output
Cs
A

$$\frac{C_f}{C_s} = \frac{R_f}{R_{in}}$$

**C_1 may be larger. If A is unity-gain compensated.

B. Compensation of stray input capacitance.


C1*** (3–10 pF typical)
R1
Rf
Input
Output
A‡
Ro‡‡
CL

$$***X_{C1} = \frac{R_1}{10} \text{ @ } f_t \text{ of A}$$

‡ A is compensated for unity gain

‡‡ $R_o \cong 50–200\ \Omega$

C. Compensation of stray output capacitance.

Figure 12-39. Stability enhancement techniques.

frequency bandwidth of 0–10 kHz, the maximum gain of the op-amp in Fig. 12-40 is 100. To have lower distortion, it would be better to have feedback at the required upper frequency limit. To increase this gain beyond 100 would require a better op-amp or two op-amps with lower gain connected in series.

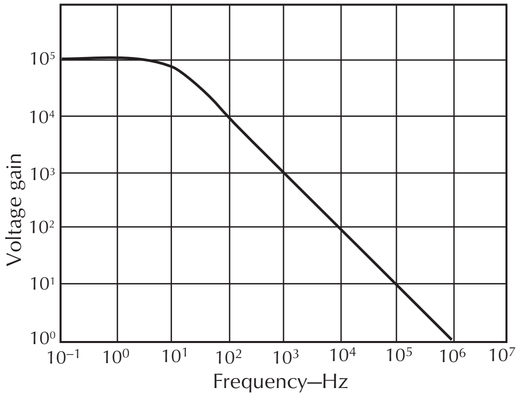
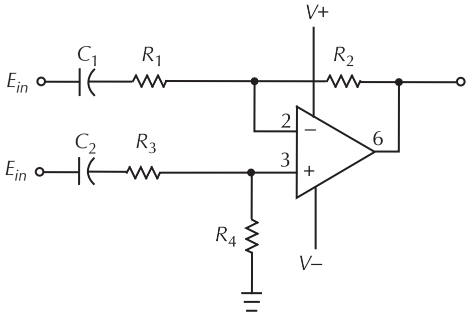
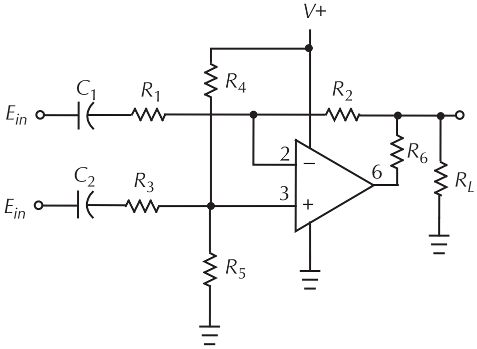


Figure 12-40. Typical open loop gain response.

Differential Amplifiers. Two *differential amplifier* circuits are shown in Fig. 12-41. The ability of the differential amplifier to block identical signals is useful to reduce hum and noise that is picked up on input lines such as in low-level microphone circuits. This rejection is called common-mode rejection and sometimes eliminates the need for an input transformer.


V+
C1
R1
R2
Ein
2
6
C2
R3
3
Ein
R4
V–

A. Basic differential amplifier.


V+
R4
C1
R1
R2
Ein
2
R6
6
C2
R3
RL
3
Ein
R5

B. Single supply differential amplifier.

Figure 12-41. Differential amplifiers.

In Fig. 12-41A, capacitors C_1 and C_2 block dc from the previous circuit and provide a 6 dB/octave rolloff below

$$f_{C_1} = \frac{1}{2\pi R_1 C_1} \tag{12-49}$$

$$f_{C_2} = \frac{1}{2\pi(R_3 + R_4)C_2} \tag{12-50}$$

The output voltage is

$$E_O = (E_{in_2} - E_{in_1})\frac{R_2}{R_1} \tag{12-51}$$

To reduce the common mode rejection ratio (CMRR),

$$\frac{R_2}{R_1} \equiv \frac{R_4}{R_3} \tag{12-52}$$

and

$$f_{C_1} = f_{C_2} \tag{12-53}$$

Summing Inverter Amplifiers. In the summing inverter, Fig. 12-32G, the virtual ground characteristic of the amplifier's summing point is used to produce a scaling adder. In this circuit, I_{in} is the algebraic sum of the number of inputs.

$$\begin{aligned} I_{in_1} &= \frac{E_{in_1}}{R_{in_1}} \\ I_{in_2} &= \frac{E_{in_2}}{R_{in_2}} \\ I_{in_n} &= \frac{E_{in_n}}{R_{in_n}} \end{aligned} \tag{12-54}$$

and the total input current is

$$\begin{aligned} I_{in} &= I_{in_1} + I_{in_2} + \ldots I_{in_n} \\ &= I_f \end{aligned} \tag{12-55}$$

and

$$I_f = \frac{-E_O}{R_f} \tag{12-56}$$

Therefore

$$I_{in_1} + I_{in_2} + \ldots\ I_{in_n} = \frac{-E_o}{R_f} \tag{12-57}$$

The output voltage is found with the equation

$$E_O = \left[R_{in_1}\left(\frac{R_f}{R_{in_1}}\right) + R_{in_2}\left(\frac{R_f}{R_{in_2}}\right) + \ldots\ R_{in_n}\left(\frac{R_f}{R_{in_n}}\right)\right] \tag{12-58}$$

It is interesting that even though the inputs mix at one point, all signals are isolated from each other and one signal does not effect the others and one impedance does not effect the rest.

Operational Transconductance Amplifiers. The *operational transconductance amplifier* (OTA) provides transconductance gain and current output rather than voltage gain and output as in an operational amplifier. The output is the product of the input voltage and amplifier transconductance, and it can be considered an infinite impedance current generator.

Varying the bias current on the OTA can completely control the open-loop gain of the device and can also control the total power input.

OTAs are useful as multipliers, automatic gain control (agc) amplifiers, sample and hold circuits, multiplexers, and multivibrators to name a few.

12.3.4 Dedicated Analog Integrated Circuits for Audio Applications

By Les Tyler and Wayne Kirkwood, THAT Corp.

The first ICs used in audio applications were general-purpose op-amps like the famous Fairchild µA741. Early op-amps like the classic 741 generally had drawbacks that limited their use in professional audio, from limited slew rate to poor clipping behavior.

Early on, IC manufacturers recognized that the relatively high-volume consumer audio market would make good use of dedicated ICs tailored to specific applications such as phono preamplifiers and companders. The National LM381 preamplifier and Signetics NE570 compander addressed the needs of consumer equipment makers producing high-volume products such as phono preamplifiers and cordless telephones. Operational Transconductance Amplifiers, such as the RCA CA3080, were introduced around 1970 to primarily serve the industrial market. It was not long before professional audio equipment manufacturers adapted OTAs for professional audio use as early voltage-controlled amplifiers (VCAs). However, through the 1970s all these integrated circuits were intended more for use in consumer and industrial applications than professional audio.

In the mid-1970s, semiconductor manufacturers began to recognize that professional audio had significantly different requirements from the needs of consumer audio or industrial products. The Philips TDA1034 was the first op-amp to combine low noise, 600 Ω drive capability and high slew rate—all important characteristics to pro audio designers. Shortly after its introduction, Philips transferred production of the TDA1034 to the newly purchased Signetics division which re-branded it the NE5534. At about the same time, Texas Instruments and National Semiconductor developed general-purpose op-amps using a combination of bipolar and FET technology (the TI TLO70- and TLO80- series, and the National LF351-series, sometimes called "BIFET"). These parts offered high slew rates, low distortion, and modest noise (though not the 600 Ω drive capability of the 5534). While not specifically aimed at pro audio, these characteristics made them attractive to pro audio designers. Along with the NE5534, these op-amps became pro audio industry standards much like the 12AX7 of the vacuum tube era.

Op-amps are fundamentally general-purpose devices. The desire to control gain via a voltage, and the application of such technology to tape noise reduction, in particular, created a market forICs that were dedicated to a specific function. This paralleled the way that phono preamplifiers spawned ICs designed for preamplification. In many ways, the VCA drove the development of early pro audio ICs.

The design of audio VCAs benefitted from the early work of Barrie Gilbert, inventor of the "Gilbert Cell" multiplier, who in 1968 published "a precise four-quadrant multiplier with subnanosecond response."[1] Gilbert discovered a current mode analog multiplication cell using current mirrors that was linear with respect to both of its inputs. Although its primary appeal at the time was to communications system designers working at RF frequencies, Gilbert laid the groundwork for many audio VCA designs.

In 1972, David E. Blackmer received U.S. Patent 3,681,618 for an "RMS Circuit with Bipolar Logarithmic Converter" and in the following year patent 3,714,462 for a "Multiplier Circuit" useful as an audio voltage-controlled amplifier. Unlike Gilbert, Blackmer used the logarithmic properties of bipolar transistors to perform the analog computation necessary for gain control and rms level detection. Blackmer's development was targeted at professional audio.[2,3]

Blackmer's timing could not have been better as the number of recording tracks expanded and, due to reduced track width coupled with the effect of summing many tracks together, tape noise increased. The expanded number of recorded tracks also increased mix complexity. Automation became a desirable feature for recording consoles because there just were not enough hands available to operate the faders.

Companies such as dbx Inc. and Dolby Laboratories benefited from this trend with tape noise reduction technologies and, in the case of dbx, VCAs for console automation. Blackmer's discrete transistor-based rms level detectors and VCAs, made by dbx, were soon used in companding multitrack tape noise reduction and console automation systems.

The early Blackmer VCAs used discrete NPN and PNP transistors that required careful selection to match each other. Blackmer's design would benefit greatly from integration into monolithic form. For some time this proved to be very difficult. Nonetheless, Blackmer's discrete audio VCAs and Gilbert's transconductance cell laid the groundwork for dedicated audio ICs. VCAs became a major focus of audio IC development.

Electronic music, not professional recording, primarily drove the early integration of monolithic VCAs and dedicated audio ICs. In 1976, Ron Dow of Solid State Music (SSM) and Dave Rossum of E-mu Systems developed some of the first monolithic ICs for analog synthesizers. SSM's first product was the SSM2000 monolithic VCA.[4] Solid State Music, later to become Solid State Microtechnology, developed an entire line of audio ICs including microphone preamplifiers, VCAs, voltage-controlled filters, oscillators, and level detectors. Later, Douglas Frey developed a VCA topology known as the o*perational voltage-controlled element* (OVCE) that was first used in the SSM2014.[5] Doug Curtis, of Interdesign and later founder of Curtis Electro Music (CEM), also developed a line of monolithic ICs for the synthesizer market that proved to be very popular with manufacturers such as Oberheim, Moog, and ARP.[6] VCAs produced for electronic music relied on NPN transistor gain cells to simplify integration.

In the professional audio market, Paul Buff of Valley People, David Baskind and Harvey Rubens of VCA Associates, and others in addition to Blackmer also advanced discrete VCA technology. Baskind and Rubens eventually produced a VCA IC that ultimately became the Aphex/VCA Associates "1537."[7]

Blackmer's VCAs and rms detectors used the precise logarithmic characteristics of bipolar transistors to perform mathematical operations suitable for VCAs and rms detection. The SSM, CEM, and Aphex products used variations on the linear multiplier, where a differential pair, or differential quad, is varied to perform VCA functions and analog voltage-controlled filtering. Close transistor matching and control of temperature-related errors are required for low distortion and control feed-through in all VCA topologies.

The Gilbert multiplier, the CA3080-series of OTAs, and the VCAs produced by SSM, CEM, and Aphex all relied solely on NPN transistors as the gain cell elements. This greatly simplified the integration of the circuits. Blackmer's log-antilog VCAs required, by contrast, precisely matched NPN and PNP transistors. This made Blackmer's VCAs the most difficult to integrate. dbx finally introduced its 2150-series monolithic VCAs in the early 1980s, almost six years after the introduction of the SSM2000.[8]

Many of the earlier developers of VCAs changed ownership or left the market as analog synthesis faded. Analog Devices currently produces many of the SSM products after numerous ownership changes. THAT Corporation assumed the patent portfolio of dbx Inc. Today Analog Devices, THAT Corporation, and Texas Instruments' Burr Brown division are the primary manufacturers making analog ICs specifically for the professional audio market.

12.3.4.1 Voltage-Controlled Amplifiers

Modern IC VCAs take advantage of the inherent and precise matching of monolithic transistors that, when combined with on-chip trimming, lowers distortion to very low levels. Two types of IC audio VCAs are commonly used and manufactured today: those based on Douglas Frey's Operational Voltage Controlled Element (OVCE)[9] and those based on David Blackmer's bipolar log-antilog topology.[10]

The Analog Devices SSM2018. The Frey OVCE gain cell was first introduced in the SSM2014 manufactured by Solid State Microtechnology (SSM).[11] SSM was acquired by Precision Monolithics, Inc, which was itself acquired by Analog Devices, who currently offers a Frey OVCE gain cell branded the SSM2018T. Frey's original patents, U.S. 4,471,320 and U.S. 4,560,947, built upon the work of David Baskind and Harvey Rubens (see U.S. Patent 4,155,047) by adding corrective feedback around the gain cell core.[12,13,14] Fig. 12-42 shows a block diagram of the SSM2018T VCA.

The OVCE is unique in that it has two outputs: V_G and $V_{1\text{-}G}$. As the V_G output increases gain with respect to control voltage, the $V_{1\text{-}G}$ output attenuates. The result is that the audio signal pans from one output to the other as the control voltage is changed.

The following expressions show how this circuit works mathematically:

$$\begin{aligned} V_{out1} &= V_G \\ &= 2K \times V_{in} \end{aligned} \tag{12-59}$$

and

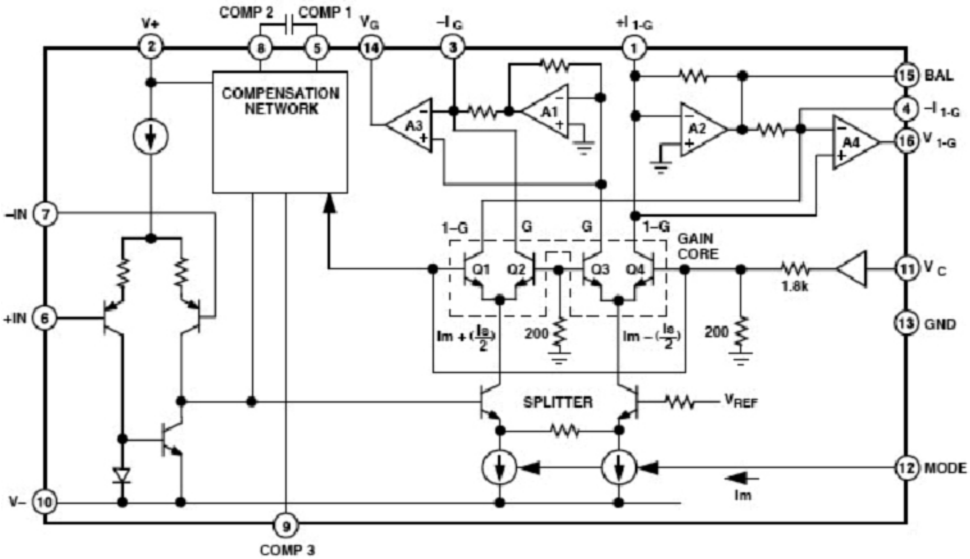


Figure 12-42. A block diagram of the SSM2018T VCA. Courtesy Analog Devices, Inc.

$$\begin{aligned} V_{out2} &= V_{1-G} \\ &= 2(1-K) \times V_{in} \end{aligned} \tag{12-60}$$

where,

K varies between 0 and 1 as the control voltage is changed from full attenuation to full gain.

When the control voltage is 0 V, $K = 0.5$ and both output voltages equal the input voltage. The value K is exponentially proportional to the applied control voltage; in the SSM2018T, the gain control constant in the basic VCA configuration is −30 mV/dB, so the decibel gain is directly proportional to the applied control voltage. This makes the part especially applicable to audio applications.

The SSM2018 has many applications as a VCA, but its use as a voltage-controlled panner (VCP) is perhaps one of the most unique, Fig. 12-43.

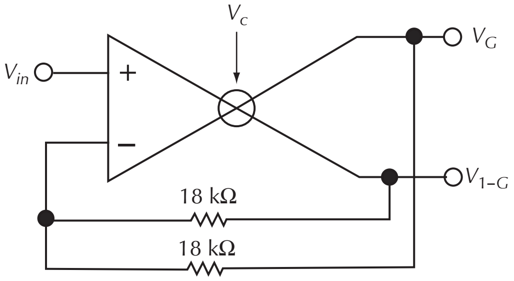


Figure 12-43. SSM2018 as a VCP. Courtesy Analog Devices, Inc.

THAT Corporation's 2180 and 2181 VCAs. The Blackmer VCAs now offered by THAT Corporation (which registered the trademark "Blackmer" for this application) exploit the mathematical property that adding a constant to the logarithm of a number is equivalent, in the linear domain, to multiplying the number by the antilog of the constant.

The equation for determining the output is

$$\begin{aligned} I_{out} &= \text{antilog}[(\log I_{in}) + E_C] \\ &= I_{in} \times [\text{antilog} E_C] \end{aligned} \tag{12-61}$$

I_{in} is multiplied by the antilog of E_C to produce I_{out}. Conveniently, and fortunately for Blackmer, the exponential response of E_C is linear in dB.

Consider the unity-gain case when $E_C = 0$.

$$\begin{aligned} I_{out} &= \text{antilog}[(\log I_{in}) + 0] \\ &= I_{in} \times [\text{antilog} 0] \\ &= I_{in} \times 1 \\ I_{out} &= I_{in} \end{aligned}$$

Blackmer VCAs exploit the logarithmic properties of a bipolar junction transistor (BJT). In the basic Blackmer circuit, the input signal I_{in} (the Blackmer VCA works in the current, not the voltage domain) is first converted to its log-domain equivalent. A control voltage, E_C, is added to the log of the input signal. Finally, the antilog is taken of the sum to provide an output signal I_{out}. This multiplies I_{in} by a control constant, E_C. When needed, the input signal voltage is converted to a current via an input resistor, and the output signal current is converted back to a voltage via an op-amp and feedback resistor.

Like the Frey OVCE, the Blackmer VCA's control voltage (E_C) is exponentiated in the process. This makes the control law exponential, or linear in dB. Many of the early embodiments of VCAs for electronic music were based on linear multiplication and required exponential converters, either external or internal to the VCA, to obtain this desirable characteristic.[15] Fig. 12-44 shows the relationship between gain and E_C for a Blackmer VCA.

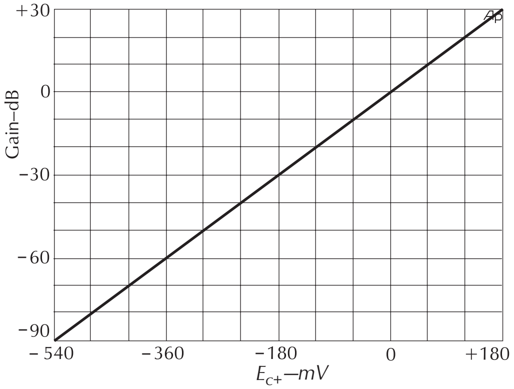


Figure 12-44. THAT 2180 gain versus E_C+. Courtesy THAT Corporation.

Audio signals are of both polarities; that is, the sign of I_{in} in the above equations will be either positive or negative at different times. Mathematically, the log of a negative number is undefined, so the circuit must be designed to handle both polarities. The essence of David Blackmer's invention was to handle each phase—positive and negative—of the signal waveform with different "genders" of transistors—NPN and PNP—and to provide a class A-B bias scheme to deal with the crossover region between the two. This made it possible to generate a sort of bipolar log and antilog. A block diagram of a Blackmer VCA is shown in Fig. 12-45.

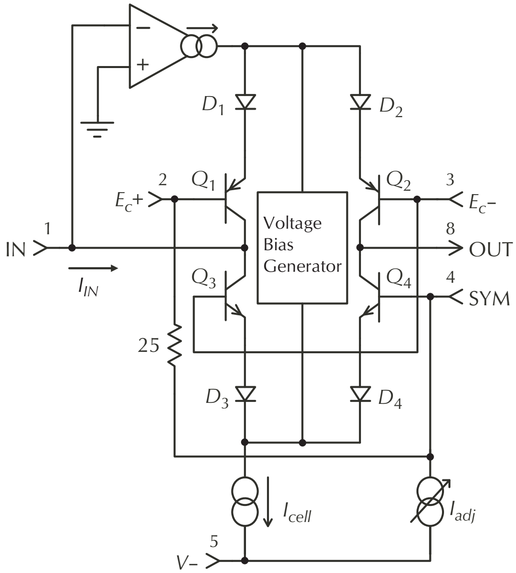


Figure 12-45. THAT 2180 equivalent schematic. Courtesy THAT Corporation.

Briefly, the circuit functions as follows. An ac input signal current I_{IN} flows in pin 1, the input pin. An internal operational transconductance amplifier (OTA) maintains pin 1 at virtual ground potential by driving the emitters of Q_1 and (through the Voltage Bias Generator) Q_3. Q_3/D_3 and Q_1/D_1 act to log the input current, producing a voltage (V_3) that represents the bipolar logarithm of the input current. (The voltage at the junction of D_1 and D_2 is the same as V_3, but shifted by four forward V_{be} drops.)

Pin 8, the output, is usually connected to a virtual ground. As a result, Q_2/D_2 and Q_4/D_4 take the bipolar antilog of V_3, creating an output current flowing to the virtual ground, which is a precise replica of the input current. If pin 2 (E_C+) and pin 3 (E_C−) are held at ground, the output current will equal the input current. For pin 2 positive or pin 3 negative, the output current will be scaled larger than the input current. For pin 2 negative or pin 3 positive, the output current is scaled smaller than the input.

The log portion of the VCA, D_1/Q_1 and D_3/Q_3, and the antilog stages, D_2/Q_2 and D_4/Q_4 in Fig. 12-45, require both the NPN and the PNP transistors to be closely matched to maintain low distortion. As well, all the devices (including the bias network) must be at the same temperature. Integration solves the matching and temperature problems, but conventional "junction-isolated" integration is notorious for offering poor-performing PNP devices. Frey and others avoided this problem by basing their designs exclusively on NPN devices for the critical multiplier stage. Blackmer's design required "good" PNPs as well as NPNs.

One way to obtain precisely matched PNP transistors that provide discrete transistor performance is to use an IC fabrication technology known as *dielectric isolation*. THAT Corporation uses dielectric isolation to fabricate integrated PNP transistors that equal or exceed the performance of NPNs. With dielectric isolation, the bottom layers of the devices are available early in the process, so both N- and P-type collectors are possible. Furthermore, each transistor is electrically insulated from the substrate and all other devices by an oxide layer, which enables discrete transistor performance with the matching and temperature characteristics only available in monolithic form.

In Fig. 12-45, it can also be seen that the Blackmer VCA has two E_C inputs having opposite control response—E_C+ and E_C−. This unique characteristic allows both control inputs to be used simultaneously. Individually, gain is exponentially proportional to the voltage at pin 2, and exponentially proportional to the negative of the voltage at pin 3. When both are used simultaneously, gain is exponentially proportional to the difference in voltage between pins 2 and 3. Overall, because of the exponential characteristic, the control voltage sets gain *linearly* in *decibels* at 6 mV/dB.

Fig. 12-46 shows a typical VCA application based on a THAT2180 IC. The audio input to the VCA is a current; an input resistor converts the input voltage to a current. The VCA output is also a current. An op-amp and its feedback resistor serve to convert the VCA's current output back to a voltage.

As with the basic topologies from Gilbert, Dow, Curtis, and other transconductance cells, the current input/output Blackmer VCA can be used as a variable conductance to tune oscillators, filters, and the like. An example of a VCA being used to control a first-order state-variable filter is shown in Fig. 12-47 with the response plot in Fig. 12-48.

When combined with audio level detectors, VCAs can be used to form a wide range of dynamics processors, including compressors, limiters, gates, duckers,

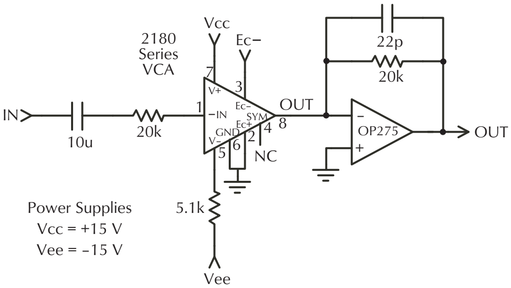


Figure 12-46. Basic THAT 2180 VCA application. Courtesy THAT Corporation.

companding noise reduction systems, and signal-controlled filters.

12.3.4.2 Peak, Average, and RMS Level Detection

It is often desirable to measure audio level for display, dynamics control, noise reduction, instrumentation, etc. Level detectors take different forms: among the most common are those that represent peak level, some form of average level over time, and *root-mean-square* (more simply known as *rms* level).

Peak signal level is usually interpreted to mean the highest instantaneous level within the audio bandwidth. Measuring peak level involves a detector with very fast charge (attack) response and much slower decay. Peak levels are often used for headroom and overload indication and in audio limiters to prevent even brief overload of transmission or storage media. However, peak measurements do not correlate well with perceived loudness, since the ear responds not only to the amplitude, but also to the duration of a sound.

Average-responding level detectors generally average out (or smooth) full or half-wave rectified signals to provide envelope information. While a pure average response (that of an R-C circuit) has equal rise (attack) and fall (decay) times, in audio applications, level detectors often have faster attack than decay. The familiar VU meter is average responding, with a response time and return time of the indicator both equal to 300 ms. The PPM meter, commonly used in Europe for audio program level measurement, combines a specific quick attack response with an equally specific, slow fall time. PPM metering provides a reliable indication of meaningful peak levels.[16]

Rms level detection is unique in that it provides an ac measurement suitable for the calculation of signal power. Rms measurements of voltage, current, or both indicate effective power. *Effective power* is the heating power of a dc signal equivalent to that offered by an ac signal. True rms measurements are not affected by the signal waveform complexity, while peak and average readings vary greatly depending on the nature of the waveform. For example, a resistor heated by a 12 Vac rms signal produces the same number of watts—and heat—as a resistor connected to 12 Vdc. This is true regardless of whether the ac waveform is a pure sinusoid, a square wave, a triangle wave or music. In instrumentation, rms is often referred to as *true rms* to distinguish it from average-responding instruments that are calibrated to read *rms* only for sinusoidal inputs. Importantly, in audio signal-processing applications, rms response is thought to closely approximate the human perception of loudness.[17]

12.3.4.3 Peak and Average Detection with Integrated Circuits

The fast response of a peak detector is often desirable for overload indication or dynamics control when a signal needs to be limited to fit the strict level confines of a transmission or storage medium. A number of op-amp-based circuits detect peak levels using full or half-wave rectification. General-purpose op-amps are quite useful for constructing peak detectors and are discussed in Section 12.3.3. The recently discontinued Analog Devices PKD01 was perhaps the only peak detector IC suited for audio applications.

Average-responding level detection is performed by rectification followed by a smoothing resistor/capacitor (R-C) filter whose time constants are chosen for the application. If the input is averaged over a sufficiently long period, the signal envelope is detected. Again, general-purpose op-amps serve quite well as rectifiers with R-C networks or integrators serving as averaging filters.

Other than meters, most simple electronic audio level detectors use an asymmetrical averaging response that attacks more quickly than it decays. Such circuits usually use diode steering to charge a capacitor quickly through a relatively small-value resistor, but discharge it through a larger resistor. The small resistor yields a fast attack, and the large resistor yields a slower decay.

12.3.4.4 Rms Level Detection Basics

Rms detection has many applications in acoustic and industrial instrumentation. As mentioned previously, rms level detectors are thought to respond similarly to

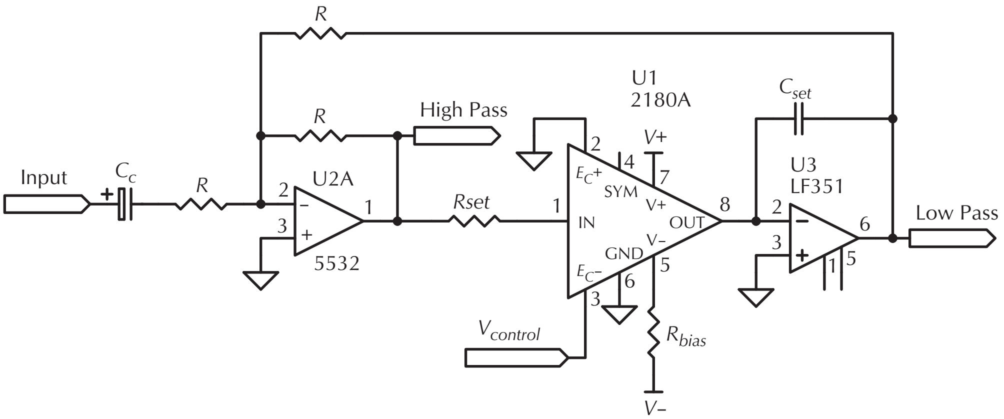


Figure 12-47. VCA state-variable filter. Courtesy THAT Corporation.

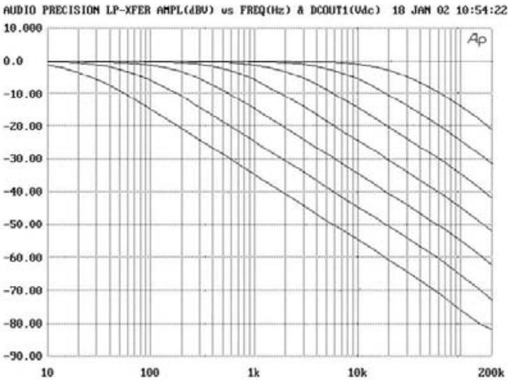


Figure 12-48. State-variable filter response. Courtesy THAT Corporation.

the human perception of loudness. This makes them particularly useful for audio dynamics control.

Rms is mathematically defined as the square root of the mean of the square of a waveform. Electronically, the mean is equal to the average, which can be approximated by an R-C network or an op-amp-based integrator. However, calculating the square and square root of waveforms is more difficult.

Designers have come up with a number of clever techniques to avoid the complexity of numerical rms calculation. For example, the heat generated by a resistive element may be used to measure power. Power is directly proportional to the square of the voltage across, or current through, a resistor, so the heat given off is proportional to the square of the applied signal level. To measure large amounts of power having very complex waveforms, such as the RF output of a television transmitter, a resistor dummy load is used to heat water. The temperature rise is proportional to the transmitter power. Such caloric instruments are naturally slow to respond, and impractical for the measurement of sound. Nonetheless, solid-state versions of this concept have been integrated, as, for example U.S. Patent 4,346,291, invented by Roy Chapel and Macit Gurol.[18] This patent, assigned to instrumentation manufacturer Fluke, describes the use of a differential amplifier to match the power dissipated in a resistive element, thus measuring the true rms component of current or voltage applied to the element. While very useful in instrumentation, this technique has not made it into audio products due to the relatively slow time constants of the heating element.

To provide faster time constants to measure small rms voltages or currents with complex waveforms such as sound, various analog computational methods have been employed. Computing the square of a signal generally requires extreme dynamic range, which limits the usefulness of direct analog methods in computing rms value. As well, the square and square-root operations require complex analog multipliers, which have traditionally been expensive to fabricate.

As with VCAs, the analog computation required for rms level detection is simplified by taking advantage of the logarithmic properties of bipolar junction transistors. The seminal work on computing rms values for audio applications was developed by David E. Blackmer, who received U.S. Patent 3,681,618 for an "RMS Circuit with Bipolar Logarithmic Converter."[17] Blackmer's circuit, discussed later, took advantage of two important log-domain properties to compute the square and square root. In the log domain, a number is squared by multiplying it by 2; the square root is obtained by dividing it by 2.

For example, to square the signal V_{in} use

$$V_{in}^{\;2} = \text{antilog}[(\log V_{in}) \times 2] \tag{12-62}$$

To take the square root of V log,

$$\sqrt{V\,\log} = \text{antilog}\left[\frac{\log(V\,\log)}{2}\right] \tag{12-63}$$

12.3.4.5 Rms Level Detection ICs

Because rms level detectors are more complex than either peak- or average-responding circuits, they benefit greatly from integration. Fortunately, a few ICs are suitable for the professional audio applications. Two ICs currently in production are the Analog Devices AD636 and the THAT Corporation THAT2252.

Analog Devices AD636. The AD636 has enjoyed wide application in audio and instrumentation. Its predecessor, the AD536, was used in the channel dynamics processor of the SSL 4000 series console in conjunction with a dbx VCA. Thousands of these channels are in daily use worldwide.

The AD636 shown in Fig. 12-49 provides both a linear-domain rms output and a dB-scaled logarithmic output. The linear output at pin 8 is ideal for applications where the rms input voltage must be read with a dc meter. Suitably scaled, 1 Vrms input can produce 1 Vdc at the buffer output, pin 6.

In audio applications such as signal processors, it is often most useful to express the signal level in dB. The AD636 also provides a dB-scaled current output at pin 5. The linear dB output is particularly useful with exponentially controlled VCAs such as the SSM2018 or THAT 2180 series.

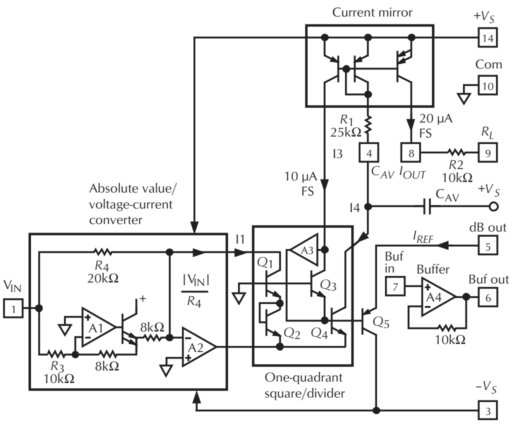


Figure 12-49. The AD636 block diagram. Courtesy Analog Devices, Inc.

Averaging required to calculate the mean of the sum of the squares is performed by a capacitor, C_{AV}, connected to pin 4. Fig. 12-50 shows an AD636 used as an audio dB meter for measurement applications.

THAT Corporation THAT2252. The 2252 IC uses the technique taught by David Blackmer to provide wide dynamic range, logarithmic linear dB output, and relatively fast time constants. Blackmer's detector delivers a fast attack with a slow linear dB decay characteristic in the log domain.[17] Because it was specifically developed for audio applications, it has become a standard for use in companding noise reduction systems and VCA-based compressor/limiters.

A simplified schematic of Blackmer's rms detector, used in the THAT2252, is shown in Fig. 12-51.

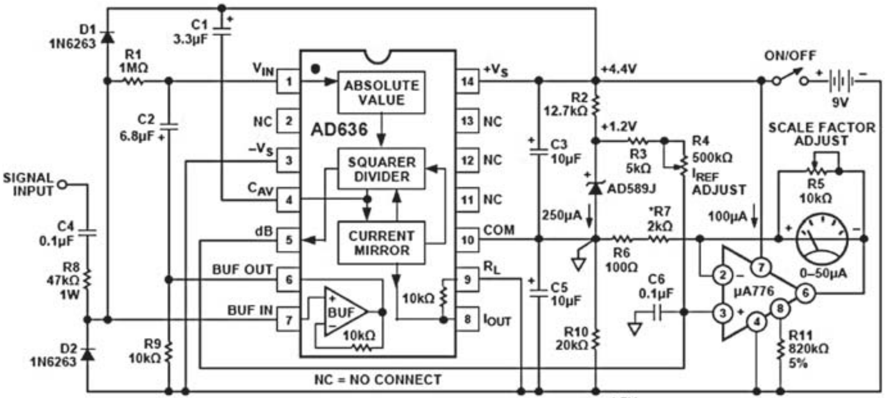
The audio input is first converted to a current I_{in} by an external resistor (not shown in Fig. 12-51). I_{in} is full-wave rectified by a current mirror rectifier formed by OA1 and Q_1-Q_3, such that IC_4 is a full-wave rectified version of I_{in}. Positive input currents are forced to flow through Q_1, and mirrored to Q_2 as IC_2; negative input currents flow through Q_3 as IC_3; both IC_2 and IC_3 thus flow through Q_4. (Note that pin 4 is normally connected to ground through an external 20 Ω resistor.)

Performing the absolute value before logarithmic conversion avoids the problem that, mathematically, the log of a negative number is undefined. This eliminates the requirement for bipolar logarithmic conversion and the PNP transistors required for log-domain VCAs.

OA2, together with Q_4 and Q_5, forms a log amplifier. Due to the two diode-connected transistors in the feedback loop of OA2, the voltage at its output is proportional to twice the log of IC_4. This voltage, $V_{\log}$, is therefore proportional to the log of I_{in}^2 (plus the bias voltage V_2).

To average $V_{\log}$, pin 6 is usually connected to a capacitor C_T and a negative current source R_T. see Fig. 12-52. The current source establishes a quiescent dc bias current, I_T, through diode-connected Q_6. Over time, C_T charges to 1 V_{be} below $V_{\log}$.

Q_6's emitter current is proportional to the antilog of its V_{be}. The potential at the base (and collector) of Q_6 represents the log of I_{in}^2 while the emitter of Q_6 is held at ac ground via the capacitor. Thus, the current in Q_6 is proportional to the square of the instantaneous change in input current. This dynamic antilogging causes the capacitor voltage to represent the log of the mean of the square of the input current. Another way to characterize the operation of Q_6, C_T, and R_T is that of a "log domain" filter.[20]


D1
1N6263
C1
3.3µF
R1
1MΩ
C2
6.8µF
SIGNAL
INPUT
C4
0.1µF
R8
47kΩ
1W
D2
1N6263
R9
10kΩ
V_IN
NC
−V_S
C_AV
dB
BUF OUT
BUF IN
ABSOLUTE
VALUE
AD636
SQUARER
DIVIDER
CURRENT
MIRROR
BUF
10kΩ
10kΩ
+V_S
NC
NC
NC
COM
R_L
I_OUT
NC = NO CONNECT
+4.4V
R2
12.7kΩ
+1.2V
R3
5kΩ
R4
500kΩ
I_REF
ADJUST
C3
10µF
AD589J
250µA
*R7
2kΩ
100µA
R6
100Ω
C6
0.1µF
C5
10µF
R10
20kΩ
µA776
ON/OFF
9V
SCALE FACTOR
ADJUST
R5
10kΩ
0–50µA
R11
820kΩ
5%
+4.7V

ALL RESISTORS 1/4W 1% METAL FILM UNLESS OTHERWISE STATED EXCEPT
*WHICH IS 2kΩ +3500ppm 1% TC RESISTOR.

Figure 12-50. AD636 as an audio dB meter. Courtesy Analog Devices, Inc.

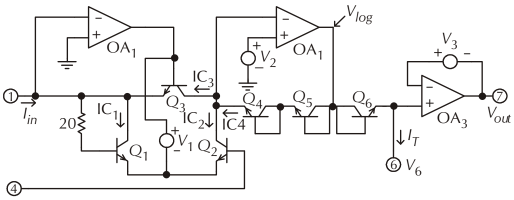

V_log
OA_1
OA_1
V_2
V_3
IC_3
Q_3
Q_4
Q_5
Q_6
①
I_in
20
IC_1
IC_2
IC4
V_1
Q_1
Q_2
I_T
⑥ V_6
OA_3
⑦
V_out
④

Figure 12-51. Block diagram of a THAT2252 IC. Courtesy THAT Corporation.

In the THAT2252, the square root portion of the rms calculation is not computed explicitly but is implied by the constant of proportionality for the output. Since, in the log domain, taking the square root is equivalent to dividing by two, the voltage at the output (pin 7) is proportional to the mean of the square at approximately 3 mV/dB and proportional to the square root of the mean of the square at approximately 6 mV/dB.

The attack and release times of rms detectors are locked in a relationship to each other and separate controls for each are not possible while still maintaining rms response. Varying the value of C_T and R_T in the THAT2252, and C_{AV} in the AD636 allow the time constant to be varied to suit the application. More complex approaches, such as a nonlinear capacitor, are possible with additional circuitry.[21]

Fig. 12-52 shows a typical application for the THAT2252. The input voltage is converted to a current by R_{in}. C_{in} blocks input dc and internal op-amp bias currents. The network around pin 4 sets the waveform symmetry for positive versus negative input currents. Internal bias for the THAT2252 is set by R_b and bypassed by a 1 µF capacitor. R_T and C_T set the timing of the log-domain filter. The output signal (pin 7) is 0 V when the input signal current equals a reference current determined by I_{bias} and I_T. It varies in dc level above and below this value to represent the dB input level at the rate of ~6 mV/dB.

Fig. 12-53 shows the tone burst response of a THAT2252, while Fig.12-54 is a plot of THAT2252 output level versus input level. The THAT2252 has linear dB response over an almost 100 dB range.

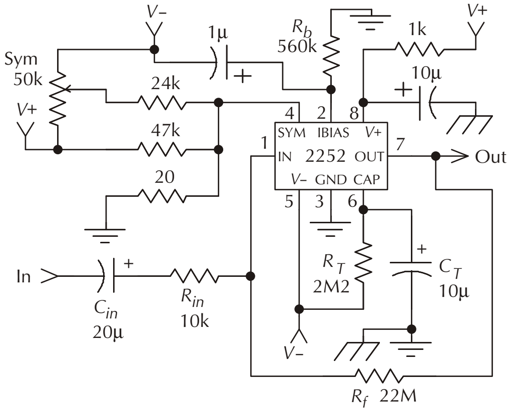

V−
1µ
R_b
560k
1k
V+
Sym
50k
24k
10µ
V+
47k
20
4
2
8
1
7
SYM IBIAS V+
IN 2252 OUT
V− GND CAP
5
3
6
Out
In
C_in
20µ
R_in
10k
R_T
2M2
C_T
10µ
V−
R_f 22M

Figure 12-52. Typical application of a THAT2252 IC. Courtesy THAT Corporation.

The Analog Devices AD636 and THAT Corporation THAT2252 provide precise, low-cost rms detection due to their integration into monolithic form. On their own, rms detectors are very useful at monitoring signal level, controlling instrumentation, and other applications. When combined with VCAs for gain control, many different signal processing functions can be realized including noise reduction, compression, and limiting.

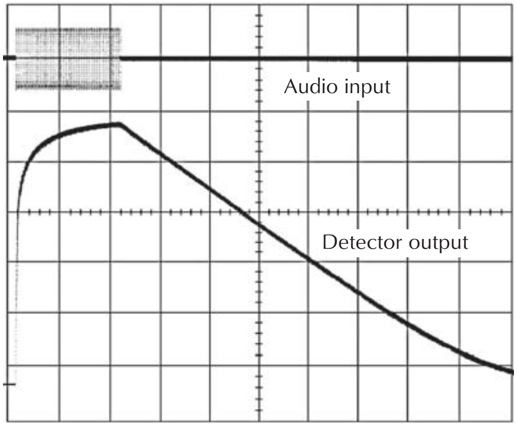


Figure 12-53. THAT2252 tone burst response. Courtesy THAT Corporation

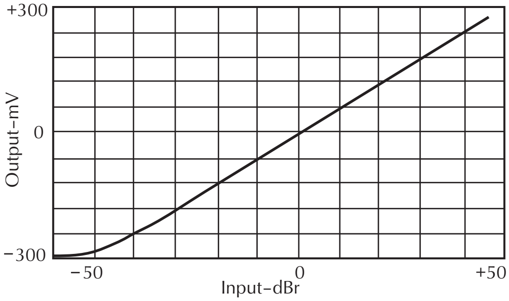


Figure 12-54. THAT2252 input versus output. Courtesy THAT Corporation.

12.3.5 Integrated Circuit Preamplifiers

The primary applications of preamplifiers for professional audio in the post-tape era are for use with microphones. Before the development of monolithic ICs dedicated to the purpose, vacuum tubes, discrete bipolar or field-effect transistors,[22] or general-purpose audio op-amps were used as preamplifiers.[23] Dynamic microphones generally produce very small signal levels and have low output impedance. Ribbon microphones are notorious for low output levels. For many audio applications, significant gain (40–60 dB) is required to bring these mic level signals up to pro audio levels. Condenser microphones, powered by phantom power, external power supplies, or batteries, often produce higher signal levels requiring less gain.

To avoid adding significant noise to the microphone's output, professional audio preamplifiers must have very low input noise. Transformer-coupled preamps ease the requirement for very low noise amplification, since they take advantage of the voltage step-up possible within the input transformer. Early transformerless, or active, designs required performance that eluded integration until the early 1980s. Until semiconductor process and design improvements permitted it and the market developed to generate sufficient demand, most microphone preamplifiers were based on discrete transistors or discrete transistors augmented with commercially available op-amps.

Virtually all professional microphones use two signal lines to produce a balanced output. This allows a preamplifier to distinguish the desired differential audio signal—which appears as a voltage *difference* between the two signal lines—from hum and noise pickup—which appears as a "common-mode" signal with the *same* amplitude and polarity on both signal lines. Common mode rejection quantifies the ability of the preamplifier to reject common mode interference while accepting differential signals.

Therefore, one goal of a pro-audio mic preamp is to amplify differential signals in the presence of common-mode hum. As well, the preamp should ideally add no more noise than the thermal noise of the source impedance—well below the self-noise of the microphone and ambient acoustic noise.

Phantom power is required for many microphones, especially professional condenser types. This is usually a +48 Vdc power supply applied to both polarities of the differential input through 6.8 kΩ resistors (one for each input polarity). Dc supply current from the microphone returns through the ground conductor. Phantom power appears in common mode essentially equal on both inputs. The voltage is used to provide power to the circuitry inside the microphone.

12.3.5.1 transformer input microphone preamplifiers

Many microphone preamplifiers use transformers at their inputs. Transformers, although costly, provide voltage gain that can ease the requirements for low noise in the subsequent amplifier. The transformer's voltage gain is determined by the turns ratio of the secondary versus the primary. This ratio also transforms impedance, making it possible to match a low-impedance microphone to a high-impedance amplifier without compromising noise performance.

A transformer's voltage gain is related to its impedance ratio by the following equation:

$$Gain = 20\log\sqrt{\frac{Z_s}{Z_p}} \tag{12-64}$$

where,
Gain is the voltage gain in dB of the transformer,
Z_p is the primary transformer impedance in ohms,
Z_s is the secondary transformer impedance in ohms.

A properly designed transformer with a 150 Ω primary and 15 kΩ secondary produces 20 dB of free voltage gain without adding noise. Well-made transformers also provide high common-mode rejection, which helps avoid hum and noise pickup. This is especially important with the low output voltages and long cable runs common with professional microphones. In addition, transformers provide galvanic isolation by electrically insulating the primary circuit from the secondary while allowing signal to pass. While usually unnecessary in microphone applications, this provides a true ground lift, which can eliminate ground loops in certain difficult circumstances.

Transformer isolation is also useful when feeding phantom power (a +48 Vdc current-limited voltage to power internal circuitry in the microphone) down the mic cable from the preamp input terminals. Phantom power may be connected through a center tap on the primary to energize the entire primary to +48 Vdc, or supplied through resistors (usually 6.8 kΩ) to each end of the primary of the transformer. (The latter connection avoids dc currents in the coils, which can lead to premature saturation of the core magnetics.) The galvanic isolation of the transformer avoids any possibility of the 48 Vdc signal from reaching the secondary windings.

12.3.5.2 Active Microphone Preamplifiers Eliminate Input Transformers

As is common in electronic design, transformers do have drawbacks. Perhaps the most prominent one is cost: a Jensen Transformer, Inc. JT-115K-E costs approximately $75 US or $3.75 per dB of gain.[24] From the point of view of signal, transformers add distortion due to core saturation. Transformer distortion has a unique sonic signature that is considered an asset or a liability—depending on the transformer and whom you ask. Transformers also limit frequency response at both ends of the audio spectrum. Furthermore, they are susceptible to picking up hum from stray electromagnetic fields.

Well-designed active *transformerless* preamplifiers can avoid these problems, lowering cost, reducing distortion, and increasing bandwidth. However, transformerless designs require far better noise performance from the active circuitry than transformer-based preamps do. Active mic preamps usually require capacitors (and other protection devices) to block potentially damaging effects of phantom power.[25]

12.3.5.3 The Evolution of Active Microphone Preamplifier ICs

Active balanced-input microphone preamplifier ICs were not developed until the early 1980s. Early IC fabrication processes did not permit high-quality low-noise devices, and semiconductor makers were uncertain of the demand for such products.

Active transformerless microphone preamplifiers must have fully differential inputs because they interface to balanced microphones. The amplifiers described here, both discrete and IC, use a current feedback CFB topology with feedback returned to one (or both) of the differential input transistor pair's emitters. Among its many attributes, current feedback permits differential gain to be set by a single resistor.

Current feedback amplifiers have a history rooted in instrumentation amplifiers. The challenges of amplifying low-level instrumentation signals are very similar to microphones. The current feedback instrumentation amplifier topology, known at least since Demrow's 1968 paper,[26] was integrated as early as 1982 as the Analog Devices AD524 developed by Scott Wurcer.[27] A simplified diagram of the AD524 is shown in Fig. 12-55. Although the AD524 was not designed as an audio preamp, the topology it used later became a de facto standard for IC microphone preamps. Demrow and Wurcer both used a bias scheme and fully balanced topology in which they wrapped op-amps around each of the two input transistors to provide both ac and dc feedback. Gain is set by a single resistor connected between the emitters (shown as 40 Ω, 404 Ω, and 4.44 kΩ), and feedback is provided by two resistors (R_{56} and R_{57}). The input stage is fully symmetrical and followed by a precision differential amplifier to convert the balanced output to single ended. Wurcer's AD524 required laser-trimmed thin film resistors with matching to 0.01% for an 80 dB common mode rejection ratio at unity gain.

Audio manufacturers, using variations on current feedback and the Demrow/Wurcer instrumentation amp, produced microphone preamps based on discrete low-noise transistor front ends as early as 1978; an example is the Harrison PC1041 module.[28] In December of 1984, Graeme Cohen also published his discrete transistor topology; it was remarkably similar to the work of Demrow, Wurcer, and the Harrison preamps.[29]

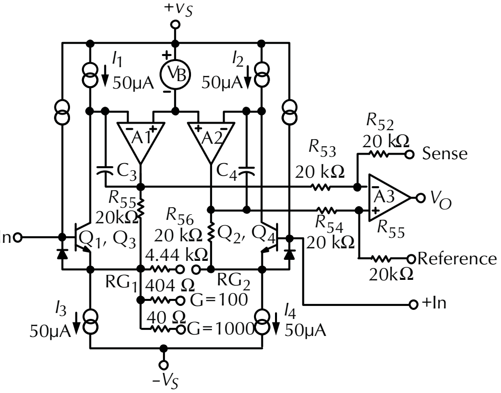


Figure 12-55. AD524 block diagram. Courtesy Analog Devices, Inc.

Solid State Music, or SSM, which later became Solid State Microtechnology, developed the first active microphone preamp IC for professional audio around 1982.[30] SSM specialized in producing niche-market semiconductors aimed at the professional audio business. The SSM2011 was almost completely self-contained, requiring only a handful of external resistors and capacitors to provide a complete preamp system. One unique feature of the SSM2011 was an on-chip LED overload and signal presence indicator.

SSM later produced the SSM2015 and the SSM2016 designed by Derek Bowers.[31] The SSM2016, and the SSM2011 and 2015 that preceded it, did not use a fully balanced topology like Wurcer's AD524 and the Harrison PC1041. The SSM parts used an internal op-amp to convert the differential stage output to single-ended. This allowed external feedback resistors to be used, eliminating the performance penalty of on-chip diffused resistors. The SSM2016 was highly successful but required external precision resistors and up to three external trims. SSM was later acquired by Precision Monolithics and eventually by Analog Devices (ADI). The SSM2016 was extremely successful and, after its discontinuance in the mid-1990s, became highly sought after.

Analog Devices introduced the SSM2017 self contained preamp, also designed by Bowers, as a replacement for the SSM2016. The SSM2017 used internal laser-trimmed thin-film resistors that permitted the fully balanced topology of the AD524 and discrete preamps to be realized as an IC. Analog Devices manufactured the SSM2017 until about 2000 when it was discontinued. A year or two later, ADI released the 2019 which is available today.

The Burr Brown division of Texas Instruments offered the INA163, which had similar performance to the SSM2017, but was not pin compatible with it. After the 2017 was discontinued, TI introduced its INA217 in the SSM2017 pinout. Today, TI produces a number of INA-family instrumentation amplifiers suitable for microphone preamps including the INA103, INA163, INA166, INA217, and the first digitally gain-controlled preamp: the PGA2500.

In 2005, THAT Corporation introduced a series of microphone preamplifiers in pinouts to match the familiar SSM2019/INA217 as well as the INA163. The THAT1510 and the performance-enhanced THAT1512 use dielectric isolation to provide higher bandwidth than the junction-isolated INA and SSM series products. (Dielectric isolation is explained in the section on audio VCAs.)

While all offer relatively high performance, the three different families of parts have different strengths and weaknesses. Differences exist in gain bandwidth, noise floor, distortion, gain structure, and supply consumption. The optimum part for any given application will depend on the exact requirements of the designer. A designer considering any one of these parts should compare their specs carefully before finalizing a new design.

12.3.5.4 Integrated Circuit Mic Preamplifier Application Circuits

The THAT1510 series block diagram is shown in Fig. 12-57. Its topology is similar to those of the TI and ADI parts. A typical application circuit is shown in Fig. 12-58. The balanced mic-level signal is applied to the input pins, In+ and In–. A single resistor (R_G), connected between pins R_{G1} and R_{G2}, sets the gain in conjunction with the internal resistors R_A and R_B. The input stage consists of two independent low-noise amplifiers in a balanced differential amplifier configuration with both ac and dc feedback returned to the emitters of the differential pair. This topology is essentially identical to the AD524 current feedback amplifier as described by Wurcer et al.

The output stage is a single op-amp differential amplifier that converts the balanced output of the gain stage into single-ended form. The THAT1500 series offers a choice of gains in this stage: 0 dB for the 1510, and –6 dB for the 1512. Gain is controlled by the input-side resistor values: 5 kΩ for the 1510 and 10 kΩ for the 1512.

The gain equations for the THAT1510 are identical to that of the SSM2017/2019 and the INA217. The INA163 and THAT 1512 have unique gain equations.

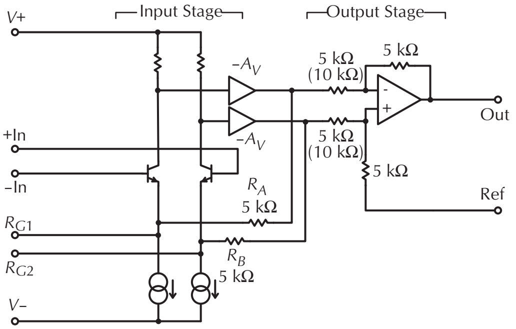


Figure 12-56. THAT1510/1512 block diagram. Courtesy THAT Corporation.

For the THAT 1510, SSM 2019, and INA217 the equation is

$$Av = 1 + \frac{10\ \text{k}\Omega}{R_G} \tag{12-65}$$

For the INA163 it is

$$Av = 1 + \frac{6\ \text{k}\Omega}{R_G} \tag{12-66}$$

For the THAT1512 it is

$$Av = 0.5 + \frac{5\ \text{k}\Omega}{R_G} \tag{12-67}$$

where,
Av is the voltage gain of the circuit.

All these parts can reach unity gain but the value of R_G required varies considerably. For the 1510, 2017, 2019, 163, and 217, gain is 0 dB ($Av = 1$) when R_G is open: this is the minimum gain of all these ICs. For the 1512, gain is –6 dB ($Av = 0.5$) with R_G open. To go from 60 dB to 0 dB gain, R_G must span a large range: 10 Ω to 10 kΩ for the 1510 and its equivalents.

R_G is typically a reverse log potentiometer (or set of switched resistors) to provide smooth rotational control of gain. In many applications, and, as shown in Fig. 12-57, a large value capacitor is placed in series with R_G to limit the dc gain of the device, thus preventing shifts in output dc-offset with gain changes. For 60 dB of gain with the THAT1512, $R_G = 5\ \Omega$ (6 Ω in the case of the INA163). Because of this, C_G must be quite large, typically ranging from 1000 µF to 6800 µF to preserve low frequency response. Fortunately, C_G does not have to support large voltages: 6.3 V is acceptable.

Parts from all manufacturers exhibit excellent voltage noise performance of $\sim 1\ \text{nV}/\sqrt{\text{Hz}}$ at high gains. Differences in noise performance begin to show up at lower gains, with the THAT 1512 offering the best performance $\sim 34\ \text{nV}/\sqrt{\text{Hz}}$ at 0 dB gain) of the group. These parts are all generally optimized for the relatively low source impedances of dynamic microphones with typically a few hundred ohm output impedance.

Fig. 12-57 provides an application example for direct connection to a dynamic microphone. Capacitors C_1–C_3 filter out radio frequencies that might cause interference (forming an RFI filter). R_1 and R_2 provide a bias current path for the inputs and terminate the microphone output. R_G sets the gain as defined in the previous equation. C_G blocks dc in the input stage feedback loop, limiting the dc gain of this stage to unity and avoiding output offset change with gain. C_6 and C_9 provide power supply bypass.

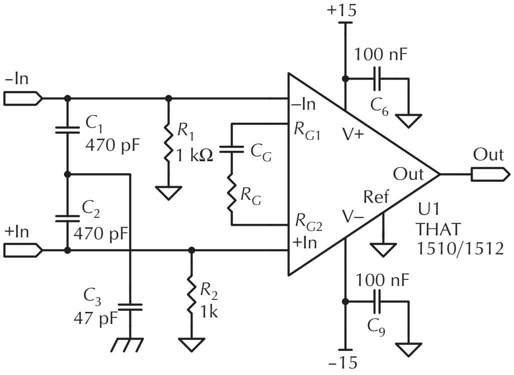


Figure 12-57. THAT1510/1512 Basic Application. Courtesy THAT Corporation.

Fig. 12-58 shows the THAT1512 used as a preamp capable of being used with phantom power. C_1–C_3 provide RFI protection. R_5 and R_6 feed phantom power to the microphone. R_9 terminates the microphone. C_4 and C_5 block 48 Vdc phantom potential from the THAT1512. R_3, R_4, and D_1–D_4 provide current limiting and overvoltage protection from phantom power faults. R_1 and R_2 are made larger than previously shown to reduce the loading on C_4 and C_5.

Many variations are possible on these basic circuits, including digital control of gain, dc servos to reduce or eliminate some of the ac-coupling needed, and exotic power supply arrangements that can produce response down to dc. For more information on possible configurations, see application notes published by Analog Devices, Texas Instruments, and THAT Corporation. (All available at their respective web sites: www.analog.com, www.ti.com, www.thatcorp.com.)

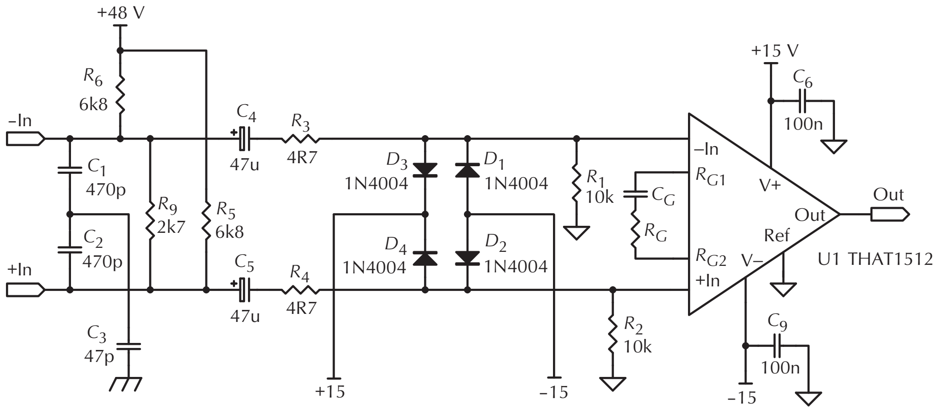


Figure 12-58. THAT preamp circuit with phantom power. Courtesy THAT Corporation.

Modern IC microphone preamplifiers provide a simple building block with performance equaling discrete solutions without a costly input transformer.

12.3.6 Balanced Line Interfaces

In professional audio, interconnections between devices frequently use *balanced lines*. These are especially important when analog audio signals are sent over long distances, where the ground references for the send and receive ends are different or where noise and interference may be picked up in the interconnection cables.

Differences in signal ground potentials arise as a result of current flowing into power-line safety grounds. These currents, flowing through finite ground impedances between equipment, can produce up to several volts potential difference between the ground references within a single building. These currents, usually at the power line frequency and its harmonics, produce the all-too-familiar hum and buzz known to every sound engineer.

Two other forms of interference, electrostatic and magnetic, also create difficulty. Cable shielding reduces electrostatic interference from fields, typically using braided copper, foil wrap, or both. Magnetic interference from fields is much harder to prevent via shielding. The impact of magnetic fields in signal cables is reduced by balanced cable construction using twisted pair cable. Balanced circuits benefit from the pair's twist by ensuring that magnetic fields cut each conductor equally. This in turn ensures that the currents produced by these fields appear in common mode, wherein the voltages produced appear equally in both inputs.

The balanced line approach comes out of telephony, in which voice communications are transmitted over many miles of unshielded twisted pair cables with reasonable fidelity and freedom from hum and interference pickup. Two principles allow balanced lines to work. First, interference—whether magnetic or electrostatic—is induced equally in both wires in the twisted paired-conductor cable, and second, the circuits formed by the source and receiver, plus the two wires connecting them form a balanced bridge,[32] Fig. 12-59. Interfering signals appear identically (in common-mode) at the two (+ and –) inputs, while the desired audio signal appears as a difference (the differential signal) between the two inputs.

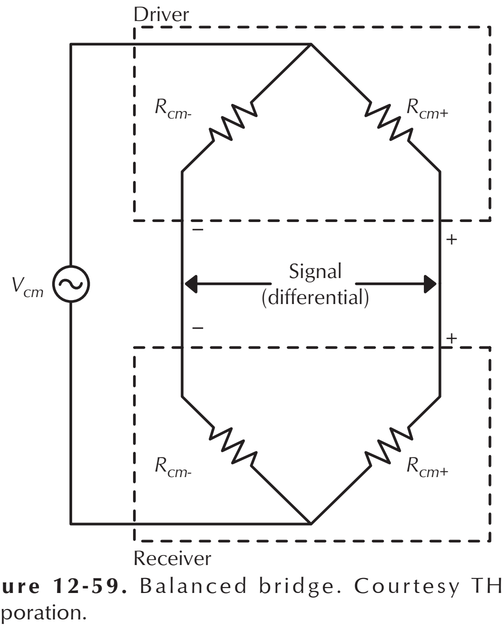


Figure 12-59. Balanced bridge. Courtesy THAT Corporation.

A common misconception in the design of balanced interfaces is that the audio signals must be transmitted as equal and opposite polarity on both lines. While this is desirable to maximize headroom in many situations, it is unnecessary to preserve fidelity and avoid noise pickup. It is enough if the bridge formed by the combination of the circuit's two common-mode source *impedances* (not the *signals*) working against the two common-mode load *impedances* remains balanced in all circumstances.

In telephony, and in early professional audio systems, transformers were used at both the inputs and outputs of audio gear to maintain bridge balance. Well-made output transformers have closely matched common-mode source impedances and very high common-mode impedance. (Common-mode impedance is the equivalent impedance from one or both conductors to ground.) The floating connections of most transformers—whether used for inputs or outputs—naturally offer very large common-mode impedance. Both of these factors, matched source impedances for output transformers, and high common-mode impedance (to ground) for both input and output transformers, work together to maintain the balance of the source/load impedance bridge across a wide range of circumstances. In addition, transformers offer *galvanic isolation,* which is sometimes helpful when faced with particularly difficult grounding situations.

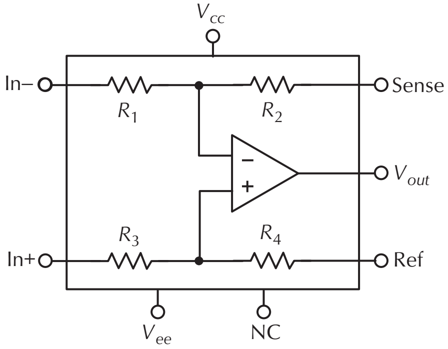
On the other hand, as noted previously in the section on preamplifiers, transformers have drawbacks of high cost, limited bandwidth, distortion at high signal levels, and magnetic pickup.

12.3.6.1 Balanced Line Inputs

Transformers were used in early balanced line input stages, particularly in the days before inexpensive op-amps made it attractive to replace them. The advent of inexpensive op-amps, especially compared to the cost of transformers, motivated the development of active transformerless inputs. As the state of the art in op-amps improved, transformer-coupled inputs were replaced by less expensive, high-performance active stages based on general-purpose parts like the Texas Instruments TL070 and TL080 series, the National Semiconductor LF351 series, and the Signetics NE5534.

As with microphone preamplifiers, common-mode rejection is an important specification for line receiver inputs. The most common configuration for active balanced line input stages used in professional audio is the simple circuit shown in Fig. 12-60. To maintain high common-mode rejection (CMR), the four resistors used must match very closely. To maintain a 90 dB CMR, for example, the resistor ratio R_1/R_2 must match that of R_3/R_4 within 0.005%. The requirement for precision-matched resistors to provide high CMR drove the development of specialized line receiver ICs.

To maintain the high CMR potential of precision balanced line receivers, the interconnections between stages must be made through low-resistance connections, and the impedances in both lines of the circuit must be very nearly identical. A few ohms of contact resistance external to the line driver and receiver (due, for example, to oxidation or poor contact) or any imbalance in the driving circuit, can significantly reduce CMR by unbalancing the bridge circuit. The imbalance can be at the source, in the middle at a cable junction, or near the input of the receiving equipment. Although many balanced line receivers provide excellent CMR under ideal conditions, few provide the performance of a transformer under less-than-ideal real world circumstances.



Part no.	Gain	R_1, R_3	R_2, R_4
THAT1240	0 dB	9 kΩ	9 kΩ
THAT1243	–3 dB	10.5 kΩ	7.5 kΩ
THAT1246	–6 dB	12 kΩ	6 kΩ

Figure 12-60. 1240 basic circuit. Courtesy THAT Corporation.

12.3.6.2 Balanced Line Outputs

Transformers were also used in early balanced output stages, for the same reasons as they are used in inputs. However, to drive 600 Ω loads, an output transformer must have more current capacity than an input transformer that supports the same voltage levels. This increased the cost of output transformers, requiring more copper and steel than input-side transformers and putting pressure on designers to find alternative outputs. Early active stages were either discrete or used discrete output transistors to boost the current available from op-amps. The NE5534, with its capability to directly drive a 600 Ω load, made it possible to use op-amps without additional buffering as output stages.

One desirable property of transformer-coupled output stages was that the output voltage was the same regardless of whether the output was connected differentially or in single-ended fashion. While professional audio gear has traditionally used balanced input stages, sound engineers commonly must interface to consumer and semi-pro gear that use single-ended input connections referenced to ground. Transformers behave just as well when one terminal of their output winding is shorted to the ground of a subsequent single-ended input stage. On the other hand, an active-balanced output stage that provides equal and opposite drive to the positive and negative outputs will likely have trouble if one output is shorted to ground.

This led to the development of a *cross-coupled* topology by Thomas Hay of MCI that allowed an active balanced output stage to mimic this property of transformers.[33] When loaded equally by reasonable impedances (e.g., 600 Ω or more) Hay's circuit delivers substantially equal—and opposite-polarity voltage signals at either output. However, because feedback is taken differentially, when one leg is shorted to ground, the feedback loop automatically produces twice the voltage at the opposing output terminal. This mimics the behavior of a transformer in the same situation.

While very clever, this circuit has at least two drawbacks. First, its resistors must be matched very precisely. A tolerance of 0.1% (or better) is often needed to ensure stability, minimize sensitivity to output loading, and maintain close matching of the voltages at either output. (Though, as noted earlier, this last requirement is unnecessary for good performance.) The second drawback is that the power supply voltage available to the two amplifiers limits the voltage swing at each output. When loaded differentially, the output stage can provide twice the voltage swing than it can when driving a single-ended load. But this means that headroom is reduced 6 dB with single-ended loads.

One way to ensure the precise matching required by Hay's circuit is to use laser-trimmed thin-film resistors in an integrated circuit. SSM was the first to do just that when they introduced the SSM2142, a balanced line output driver with a cross-coupled topology.

12.3.6.3 Integrated Circuits for Balanced Line Interfaces

Instrumentation amplifier inputs have similar requirements to those of an audio line receiver. The INA105, originally produced by Burr Brown and now Texas Instruments, was an early instrumentation amplifier that featured laser-trimmed resistors to provide 86 dB common-mode rejection. Although its application in professional audio was limited due to the performance of its internal op-amps, the INA105 served as the basis for the modern audio-balanced line receiver.

In 1989, the SSM Audio Products Division of Precision Monolithics introduced the SSM2141 balanced line receiver and companion SSM2142 line driver. The SSM2141 was offered in the same pinout as the INA105 but provided low noise and a slew rate of almost 10 V/µs. With a typical CMR of 90 dB, the pro-audio industry finally had a low-cost, high-performance replacement for the line input transformer. The SSM2142 line driver, with its cross-coupled outputs, became a low-cost replacement for the output transformer. Both parts have been quite successful.

Today, Analog Devices (which acquired Precision Monolithics) makes the SSM2141 line receiver and the SSM2142 line driver. The SSM2143 line receiver, designed for 6 dB attenuation, was introduced later to offer increased input headroom. It also provides overall unity gain operation when used with an SSM2142 line driver, which has 6 dB of gain.

The Burr Brown division of Texas Instruments now produces a similar family of balanced line drivers and receivers, including dual units. The INA134 audio differential line receiver is a second source to the SSM2141. The INA137 is similar to the SSM2143 and also permits gains of ±6 dB. Both devices owe their pinouts to the original INA105. Dual versions of both parts are available as the INA2134 and 2137. TI also makes cross-coupled line drivers known as the DRV134 and DRV135.

THAT Corporation also makes balanced line drivers and receivers. THAT's 1240 series single and 1280 series dual balanced line receivers use laser-trimmed resistors to provide high common rejection in the familiar SSM2141 (single) and INA2134 (dual) pinouts. For lower cost applications, THAT offers the 1250- and 1290-series single and dual line receivers. These parts eliminate laser trimming, which sacrifices CMR to reduce cost. Notably, THAT offers both dual and single line receivers in the unique configuration of ±3 dB gain, which can optimize dynamic range for many common applications.

THAT Corporation also offers a unique line receiver, the THAT1200 series, based on technology licensed from William E. Whitlock of Jensen Transformers, Inc. (U.S. Patent 5,568,561).[34] This design, dubbed InGenius (a trademark of THAT Corporation), bootstraps the common-mode input impedance to raise it into the megohm range of transformers. This overcomes the loss of common-mode rejection when the impedances feeding the line receiver are slightly unbalanced and permits transformer like operation. The InGenius circuit will be discussed in a following section.

THAT also offers the THAT1646 balanced line driver, which has identical pinout to the SSM2142 and DRV134/135. THAT's 1606 balanced line driver is unique among these parts in that it provides not only a differential output, but also a differential input—enabling a more direct connection to digital to analog converters.

The THAT1646 and 1606 use a unique output topology unlike conventional cross-coupled outputs which THAT calls "OutSmarts" (another trademark). OutSmarts is based on U.S. Patent 4,979,218 issued to Chris Strahm, then of Audio Teknology Incorporated.[35] Conventional cross-coupled outputs lose common-mode feedback when one output is shorted to ground to accommodate a single-ended load. This allows large signal currents to flow into ground, increasing crosstalk and distortion. Strahm's circuit avoids this by using an additional feedback loop to provide current feedback. Application circuits for the THAT1646 will be described in the section "Balanced Line Outputs."

12.3.6.4 Balanced Line Input Application Circuits

Conventional balanced line receivers from Analog Devices, Texas Instruments, and THAT Corporation are substantially equivalent to the THAT1240 circuit shown in Fig. 12-61. Some variations exist in the values of R_1–R_4 from one manufacturer to the other that will influence input impedance and noise. The ratio of R_1/R_3 to R_2/R_4 establishes the gain with $R_1 = R_3$ and $R_2 = R_4$. V_{out} is normally connected to the sense input resistor with the reference pin grounded.

Line receivers usually operate at either unity gain (SSM2141, INA134, THAT1240, or THAT1250) or in attenuation (SSM2143, INA137, THAT1243, or THAT1246, etc.). When a perfectly balanced signal (with each input line swinging ½ the differential voltage) is converted from differential to single-ended by a unity gain receiver, the output must swing twice the voltage of either input line for a net voltage gain of +6 dB. With only +21 dBu output voltage available from a line receiver powered by bipolar 15 V supplies, additional attenuation is often needed to provide headroom to accommodate pro audio signal levels of +24 dBu or more. The ratios R_1/R_2 and R_3/R_4 are 2:1 in the SSM2143, INA137, and THAT1246 to provide 6 dB attenuation. These parts accommodate up to +27 dBu inputs without clipping their outputs when running from bipolar 15 V supplies. The THAT1243, and THAT's other ±3 dB parts (the 1253, 1283, and 1293) are unique with their 0.707 attenuation. This permits a line receiver that accommodates +24 dBu inputs but avoids additional attenuation that increases noise. A –3 dB line receiver is shown in Fig. 12-62.

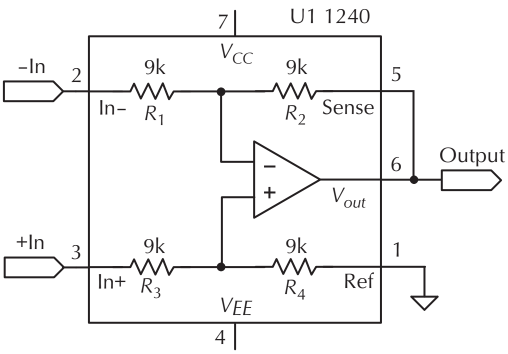


Figure 12-61. THAT1240 with 0 dB gain. Courtesy THAT Corporation.

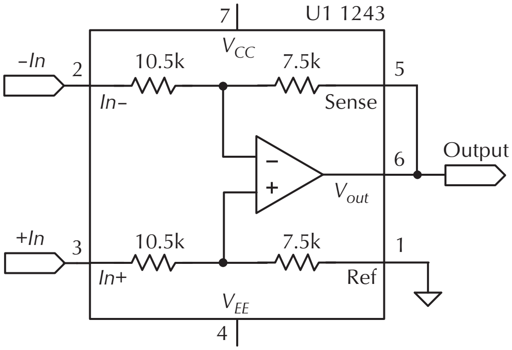


Figure 12-62. THAT1243 with 3 dB attenuation. Courtesy THAT Corporation.

The ±6 dB parts from all three manufacturers (and the ±3 dB parts from THAT) may be configured for gain instead of attenuation. To accomplish this, the reference and sense pins are used as inputs with the *In*– pin connected to V_{out} and the *In*+ pin connected to ground. A line receiver configured for 6 dB gain is shown in Fig. 12-63.

Balanced line receivers may also be used to provide sum-difference networks for mid-side (M/S or M-S) encoding/decoding as well as general-purpose applications requiring precise difference amplifiers. Such applications take advantage of the precise matching of resistor ratios possible via monolithic, laser-trimmed

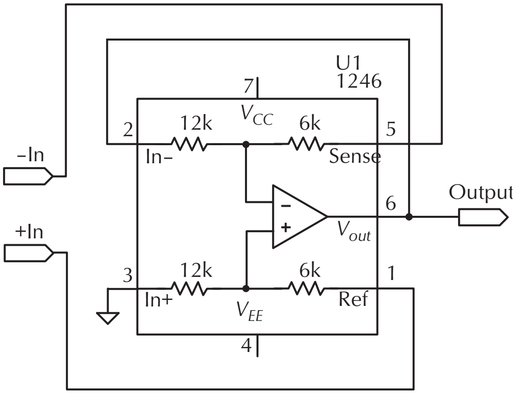


Figure 12-63. THAT 1246 with 6 dB gain. Courtesy THAT Corporation.

resistors. In fact, while these parts are usually promoted as input stages, they have applications to many circuits where precise resistor ratios are required. The typical 90 dB common-mode rejection advertised by many of these manufacturers requires ratio matching to within 0.005%.

Any resistance external to the line receiver input appears in series with the highly matched internal resistors. A basic line receiver connected to an imbalanced circuit is shown in Fig. 12-64. Even a slight imbalance, one as low as 10 Ω from connector oxidation or poor contact, can degrade common-mode rejection. Fig. 12-65 compares the reduction in CMR for low common-mode impedance line receivers versus the THAT1200 series or a transformer.

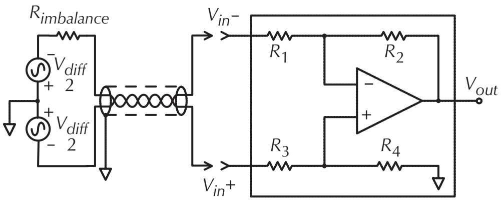


Figure 12-64. Balanced circuit with imbalance. Courtesy THAT Corporation.

The degradation of common-mode rejection from impedance imbalance comes from the relatively low-impedance load of simple line receivers interacting with external impedance imbalances. Since unwanted hum and noise appear in common-mode (as the same signal in both inputs), common-mode loading by common-mode input impedance is often a significant source of error. (The differential input impedance is the load seen by differential signals; the common-mode input impedances is the load seen by common-mode signals.) To reduce the effect of impedance imbalance, the common-mode input impedance, but not the differential impedance, must be made very high.

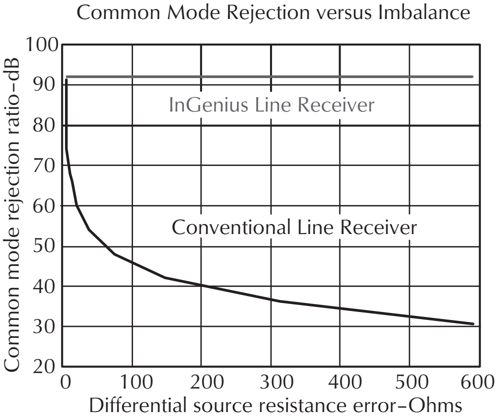


Figure 12-65. CMR imbalance versus source. Courtesy THAT Corporation.

12.3.6.5 Balanced Line Receivers with the Common-Mode Performance of a Transformer

The transformer input stage has one major advantage over most active input stages: its common-mode input impedance is extremely high regardless of its differential input impedance. This is because transformers offer floating connections without any connection to ground. Active stages, especially those made with the simple SSM2141-type IC have common-mode input impedances of approximately the same value as their differential input impedance. (Note that for simple differential stages such as these, the common-mode and differential input impedances are not always the same.) Op-amp input bias current considerations generally make it difficult to use very high impedances for these simple stages. A bigger problem is that the noise of these stages increases with the square root of the impedances chosen, so large input impedances inevitably cause higher noise.

Noise and op-amp requirements led designers to choose relatively low impedances (10 k~25 kΩ). Unfortunately, this means these stages have relatively low common-mode input impedance as well (20 k~50 kΩ). This interacts with the common-mode output impedance (also relative to ground) of the driving stage, and added cable or connector resistance. If the driver, cable, or connectors provide an unequal, nonzero common-mode output impedance, the input stage loading will upset the natural balance of any common-mode signal, converting it from common-mode to differential. No amount of precision in the input stage's resistors will reject this common-mode-turned-to-differential signal. This can completely spoil the apparently fine performance available from the precisely matched resistors in simple input stages.

An instrumentation amplifier, Fig. 12-66, may be used to increase common-mode input impedance. Input resistors R_{i1} and R_{i2} must be present to supply a bias current return path for buffer amplifiers OA1 and OA2. R_{i1} and R_{i2} can be made large—in the MΩ range—to minimize the effect of impedance imbalance. While it is possible to use this technique to make line receivers with very high common-mode input impedances, doing so requires specialized op-amps with bias-current compensation or FET input stages. In addition, this requires two more op-amps in addition to the basic differential stage (OA3).

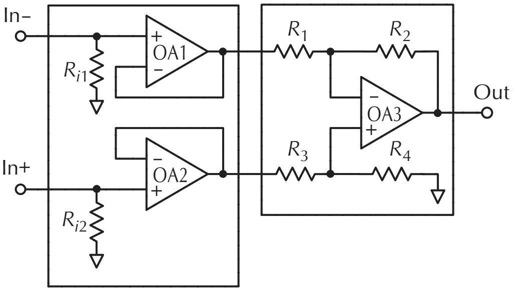


Figure 12-66. Instrumentation amplifier. Courtesy THAT Corporation.

With additional circuitry, even higher performance can be obtained by modifying the basic instrumentation amplifier circuit. Bill Whitlock of Jensen Transformers, Inc.developed and patented (U.S. Patent 5,568,561) a method of applying bootstrapping to the instrumentation amplifier in order to further raise common-mode input impedance.[34] THAT Corporation incorporated this technology in its InGenius series of input stage ICs.

12.3.6.6 InGenius High Common-Mode Rejection Line Receiver ICs

Fig. 12-67 shows the general principle behind ac bootstrapping in a single-ended connection. By feeding the ac component of the input into the junction of R_a and R_b, the effective value of R_a (at ac) can be made to appear quite large. The dc value of the input impedance (neglecting R_s being in parallel) is $R_a + R_b$. Because of bootstrapping, R_a and R_b can be made relatively small values to provide op-amp bias current, but the ac load on R_s (Z_{in}) can be made to appear to be extremely large relative to the actual value of R_a.

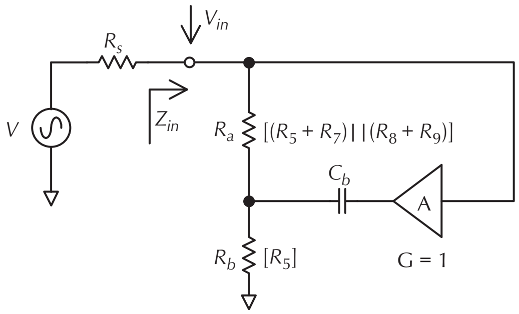


Figure 12-67. Single ended bootstrap. Courtesy THAT Corporation.

A circuit diagram of an InGenius balanced line receiver using the THAT1200 is shown in Fig. 12-68. (All the op-amps and resistors are internal to the IC.) R_5–R_9 provides dc bias to internal op-amps OA1 and OA2. Op-amp OA4, along with R_{10} and R_{11} extract the common-mode component at the input and feed the ac common-mode component back through C_b to the junction of R_7 and R_8. Because of this positive feedback, the effective values of R_7 and R_8—at ac—are multiplied into the MΩ range. In its data sheet for the 1200 series ICs, THAT cautions that C_b should be at least 10 µf to maintain common-mode input impedance (Z_{inCM}) of at least 1 MΩ at 50 Hz. Larger capacitors can increase Z_{inCM} at low power-line frequencies up to the IC's practical limit of ~10 MΩ. This limitation is due to the precision of the gain of the internal amplifiers.

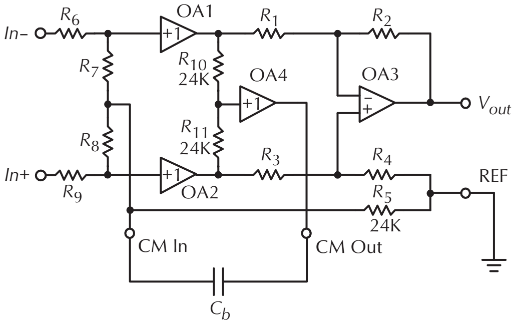


Figure 12-68. Balanced line receiver. Courtesy THAT Corporation.

The outputs of OA1 and OA2 contain replicas of the positive and negative input signals. These are converted to single-ended form by a precision differential amplifier OA3 and laser-trimmed resistors R_1–R_4. Because OA1 and OA2 isolate the differential amplifier, and the positive common-mode feedback ensures very high common-mode input impedance, a 1200-series input stage provides 90 dB CMR even with high levels of imbalance.

It took Bill Whitlock and Jensen Transformers, Inc. to provide an active input as good as a transformer operating under conditions likely to be found in the real world.

A basic application circuit using the THAT1200 series parts is shown in Fig. 12-69.

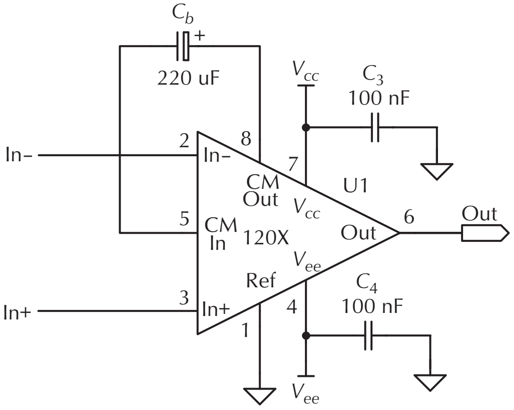


Figure 12-69. InGenius basic application. Courtesy THAT Corporation.

12.3.6.7 Balanced Line Drivers

The Analog Devices SSM2142 and Texas Instruments DRV series balanced line drivers use a cross-coupled method to emulate a transformer's floating connection and provide constant level with both single-ended (grounded) terminations and fully balanced loads. A block diagram of a cross-coupled line driver is shown in Fig. 12-70. The force and sense lines are normally connected to each output either directly or through small electrolytic coupling capacitors. A typical application of the SSM2142 driving an SSM2141 (or SSM2143) line receiver is provided in Fig. 12-71.

If one output of the cross-coupled line driver outputs is shorted to ground in order to provide a single-ended termination, the full short-circuit current of the device will flow into ground. Although this is not harmful to the device, and is in fact a recommended practice, large clipped signal currents will flow into ground, which can produce crosstalk within the product using the stage, as well as in the output signal line itself.

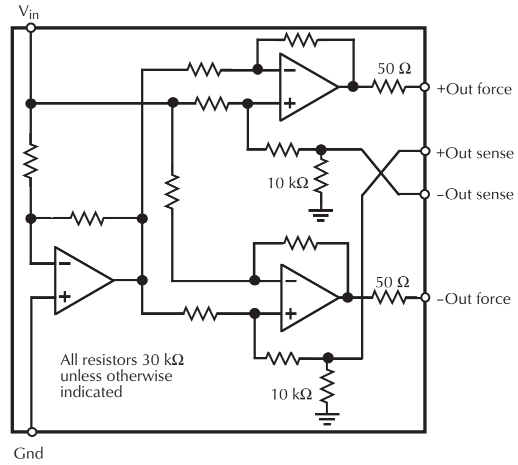


Figure 12-70. SSM2142 cross coupled output. Courtesy Analog Devices, Inc.

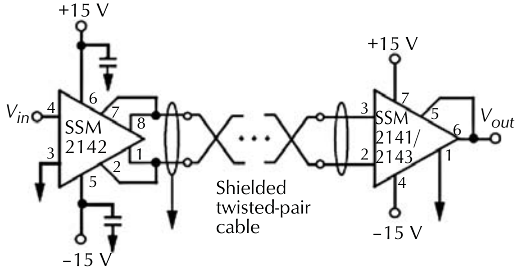


Figure 12-71. SSM2142 driving a SSM2141 line receiver. Courtesy Analog Devices, Inc.

THAT Corporation licensed a patented technology developed by Chris Strahm of Audio Teknology Incorporated. U.S. Patent 4,979,218, issued in December 1990, describes a balanced line driver that emulates a floating transformer output by providing a current-feedback system where the current from each output is equal and out of phase to the opposing output.[35] THAT trademarked this technology as OutSmarts and introduced its THAT1646 line driver having identical pinout and functionality to the SSM2142. THAT also offers a version of the 1646 with differential inputs known as the THAT1606. Fig. 12-72 is a simplified block diagram of the THAT1646.

The THAT1646 OutSmarts internal circuitry differs from other manufacturer's offerings. Outputs D_{out-} and D_{out+} supply current through 25 Ω build-out resistors. Feedback from both sides of these resistors is returned into *two* internal common-mode feedback paths. The driven side of the build-out resistors are fed back into the common-mode C_{in-} input while the load side of the build-out resistors, through the sense– and sense+ pins,

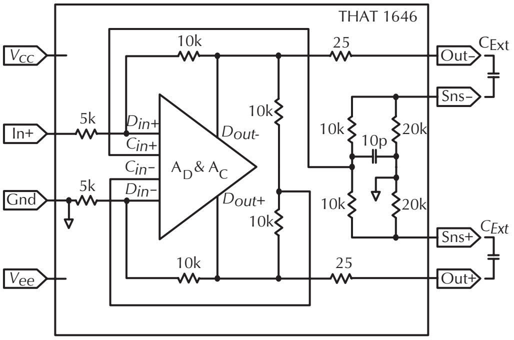

THAT 1646
Vcc
In+
Gnd
Vee
5k
5k
10k
10k
10k
10k
25
25
Din+
Cin+
Cin–
Din–
AD & AC
Dout-
Dout+
10k
10p
20k
10k
20k
Out–
Sns–
Sns+
Out+
CExt
CExt

Figure 12-72. THAT 1646 block diagram. Courtesy THAT Corporation.

provide feedback into the C_{in+} input. A current feedback bridge circuit allows the 1646 to drive one output shorted to ground to allow a single-ended load to be connected. The output short increases gain by 6 dB, similarly to conventional cross-coupled topologies. However, it does so without loss of the common-mode feedback loop. The resulting current feedback prevents large, clipped signal currents flowing into ground. This reduces the crosstalk and distortion produced by these currents.

A typical application circuit for the THAT1646 is shown in Fig. 12-73.

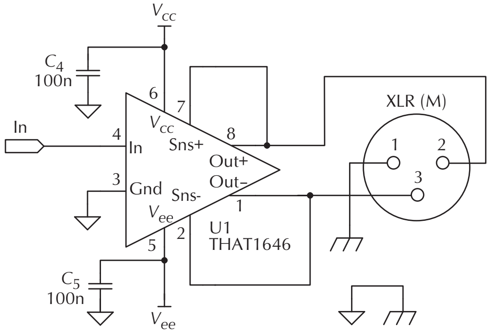

Vcc
C4
100n
In
6
7
4
3
Vcc
In
Sns+
8
Out+
Out–
Gnd
Sns-
1
Vee
2
5
U1
THAT1646
XLR (M)
1
2
3
C5
100n
Vee

Figure 12-73. THAT 1646 application. Courtesy THAT Corporation.

To reduce the amount of common-mode dc offset, the circuit in Fig. 12-74 is recommended. Capacitors C_1 and C_2, outside the primary signal path, minimize common-mode dc gain, which reduces common-mode output offset voltage and the effect of OutSmarts at low frequencies. Similar capacitors are used in the ADI and TI parts to the same effect, although OutSmarts' current feedback does not apply.

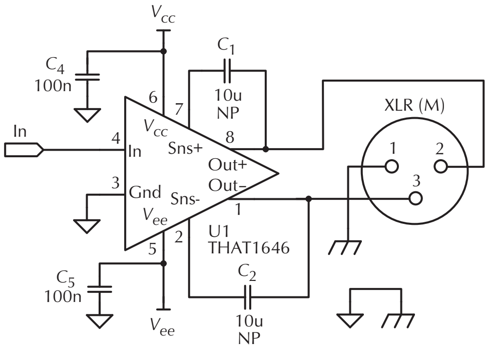

Vcc
C4
100n
C1
10u
NP
In
6
7
4
Vcc
In
Sns+
8
Out+
Out–
3
Gnd
Sns-
1
Vee
2
5
U1
THAT1646
XLR (M)
1
2
3
C5
100n
C2
10u
NP
Vee

Figure 12-74. THAT 1646 CMR offset reduction circuit. Courtesy THAT Corporation.

THAT's 1606 version of OutSmarts provides a differential input for easier connection to a digital-to-analog converter's output. A typical application of the THAT1606 is shown in Fig. 12-75. Another advantage to the 1606 is that it requires only a single low-value capacitor (typically a film type) versus the two larger capacitors required by the THAT1646, SSM2142, or DRV134.

Active balanced line drivers and receivers offer numerous advantages over transformers providing lower cost, weight, and distortion, along with greater bandwidth and freedom from magnetic pickup. When used properly, active devices perform as well as, and in many ways better than, the transformers they replace. With careful selection of modernIC building blocks from several IC makers, excellent performance is easy to achieve.

12.3.7 Digital Integrated Circuits

Digital ICs produce an output of either 0 or 1. With digital circuits, when the input reaches a preset level, the output switches polarity. This makes digital circuitry relatively immune to noise.

Bipolar technology is characterized by very fast propagation time and high power consumption, while MOS technology has relatively slow propagation times, low power consumption, and high circuit density. Fig. 12-76 shows typical circuits and characteristics of the major bipolar logic families.

Table 12-4 gives some of the terminology common to digital circuitry and digital ICs.

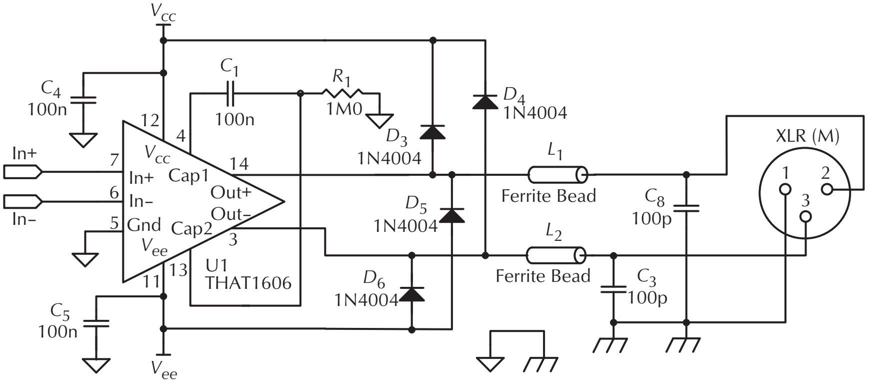


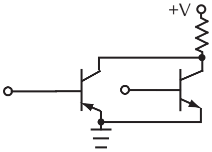
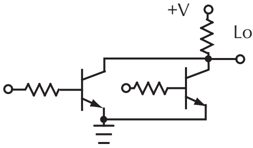
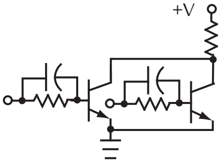
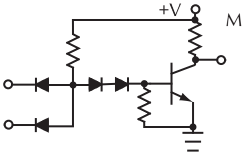
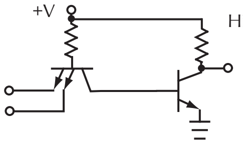
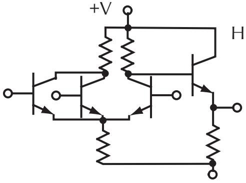
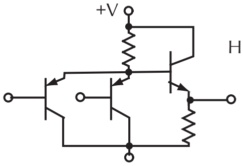
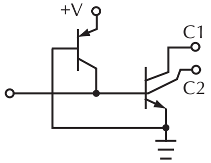
Figure 12-75. THAT 1606 application. Courtesy THAT Corporation.

Table 12-4. Digital Circuit Terminology

Adder	Switching circuits that generate sum and carry bits.
Address	A code that designates the location of information and instructions.
AND	A Boolean logic operation that performs multiplication. All inputs must be true for the output to be true.
Asynchronous	A free-running switching network that triggers successive instructions.
Bit	Abbreviation for binary digit; a unit of binary information.
Buffer	A noninverting circuit used to handle fan-out or convert input and output levels.
Byte	A fixed-length binary-bit pattern (word).
Clear	To restore a device to its standard state.
Clock	A pulse generator used to control timing of switching and memory circuits.
Clock rate	The frequency (speed) at which the clock operates. This is normally the major speed of the computer.
Counter	A device capable of changing states in a specified sequence or number of input signals.
Counter, binary	A single input flip-flop. Whenever a pulse appears at the input, the flip-flop changes state (called a *T flip-flop*).
Counter, ring	A loop or circuit of interconnected flip-flops connected so that only one is on at any given time. As input signals are received, the position of the on state moves in sequence from one flip-flop to another around the loop.
Fan-in	The number of inputs available on a gate.
Fan-out	The number of gates that a given gate can drive. The term is applicable only within a given logic family.
Flip-flop	A circuit having two stable states and the ability to change from one state to the other on application of a signal in a specified manner.
Flip-flop, D	*D* stands for delay. A flip-flop whose output is a function of the input that appeared one pulse earlier; that is, if a 1 appears at its input, the output will be a 1 a pulse later.
Flip-flop, JK	A flip-flop having two inputs designated *J* and *K*. At the application of a clock pulse, a 1 on the *J* input will set the flip-flop to the 1 or on state; a 1 on the K input will reset it to the 0 or off state; and 1s simultaneously on both inputs will cause it to change state regardless of the state it had been in.
Flip-flop, RS	A flip-flop having two inputs designated *R* and *S*. At the application of a clock pulse, a 1 on the S input will set the flip-flop to the 1 or on state, and a 1 on the *R* input will reset it to the 0 or off state. It is assumed that 1s will never appear simultaneously at both inputs.
Flip-flop, R, S, T	A flip-flop having three inputs, *R*, *S*, and *T*. The *R* and *S* inputs produce states as described for the *RS* flip-flop above; the *T* input causes the flip-flop to change states.
Flip-flop, T	A flip-flop having only one input. A pulse appearing on the input will cause the flip-flop to change states.
Gate	A circuit having two or more inputs and one output, the output depending on the combination of logic signals at the inputs. There are four gates: AND, OR, NAND, NOR. The definitions below assume positive logic is used.
Gate, AND	All inputs must have 1-state signals to produce a 0-state output.
Gate, NAND	All inputs must have 1-state signals to produce a 1-state output.
Gate, NOR	Any one or more inputs having a 1-state signal will yield a 0-state output.
Gate, OR	Any one or more inputs having a 1-state signal is sufficient to produce a 1-state output.

Table 12-4. Digital Circuit Terminology (Continued)

Inverter	The output is always in the opposite logic state as the input. Also called a NOT circuit.
Memory	A storage device into which information can be inserted and held for use at a later time.
NAND gate (D = ABC for positive inputs)	The simultaneous presence of all inputs in the positive state generates an inverted output.
Negative logic	The more negative voltage (or current) level represents the 1-state; the less negative level represents the 0-state.
NOR gate (D = A + B + C for positive inputs)	The presence of one or more positive inputs generates an inverted output.
NOT	A boolean logic operator indicating negation. A variable designated NOT will be the opposite of its AND or OR function. A switching function for only one variable.
OR	A boolean operator analogous to addition (except that two truths will only add up to one truth). Of two variables, only one need be true for the output to be true.
Parallel operator	Pertaining to the manipulation of information within computer circuits in which the digits of a word are transmitted simultaneously on separate lines. It is faster than serial operation but requires more equipment.
Positive logic	The more positive voltage (or current) level represents the 1-state; the less positive level represents the 0-state.
Propagation delay	A measure of the time required for a change in logic level to spread through a chain of circuit elements.
Pulse	A change of voltage or current of some finite duration and magnitude. The duration is called the *pulse width* or *pulse length*; the magnitude of the change is called the *pulse amplitude* or *pulse height*.
Register	A device used to store a certain number of digits in the computer circuits, often one word. Certain registers may also include provisions for shifting, circulating, or other operations.
Rise time	A measure of the time required for a circuit to change its output from a low level (zero) to a high level (one).
Serial operation	The handling of information within computer circuits in which the digits of a word are transmitted one at a time along a single line. Though slower than parallel operation, its circuits are much less complex.
Shift register	An element in the digital family that uses flip-flops to perform a displacement or movement of a set of digits one or more places to the right or left. If the digits are those of a numerical expression, a shift may be the equivalent of multiplying the number by a power of the base.
Skew	Time delay or offset between any two signals.
Synchronous timing	Operation of a switching network by a clock pulse generator. Slower and more critical than asynchronous timing but requires fewer and simpler circuits.
Word	An assemblage of bits considered as an entity in a computer.

Symbol	Circuit Diagram	Speed*	Power*	Fan-Out*	Noise Immunity*	Trade Name	Remarks
DCTL		Medium	Medium	Low	Low	Series 53	Variations in input characteristics result in base current "hogging" problems. Proper operation not always guaranteed. More susceptibile to noise because of low operating and signal voltages.
RTL		Low	Low	Low	Low	RTL	Very similar to DCTL. Resistors resolve current "hogging" problem and reduce power dissipation. However,operating speed is reduced.
RCTI		Low	Low	Low	Low	Series 51	Though capacitors can increase speed capability, noise immunity is affected by capacitive coupling of noise signals.
DTL		Medium	Medium	Medium	Medium to high	930 DTL	Use of pull-up resistor and charge-control technique improves speed capabilities. Many variations of this circuit exist, each having specific advantages.
TTL		High	Medium	Medium	Medium to high	SUHL Series 54/74	Very similar to DTL. Has lower parasitic capacity at inputs. With the many existing variations, this has become very
CML (ECL)		High	High	High	Medium to high	MECL ECCSL	Similar to a differential amplifier, the reference voltage sets the threshold voltage. High-speed, high-fan-out operation is possible with associated high power dissipation. Also known as *emitter-coupled logic* (ECL).
CTL		High	High	Medium	Medium	CTML	More difficult manufacturing process results in compromises of active device characteristics and higher cost.
I^2L		High	Low	High	Medium	I^2L	Provides smallest and most dense bipolar gate. Simple manufacturing process and higher component packing density than the MOS process. Also known as merged-transistor logic (MTL).

*Low =	<5 MHz	<5 mW	<5	<300 mV
Medium =	5–15 MHz	5–15 mW	5–10	300–500 mV
High =	>15 MHz	>15mW	>10	>500 mV

Figure 12-76. Typical digital circuits and their characteristics for the major logic families. (Adapted from Reference 4.)

References

1. Gilbert, “A precise Four Quadrant Multiplier with Subnanosecond Response,” IEEE *J. of Solid State Circuits*, Vol. SC-3, no. 4, December 1968.
2. Blackmer, “RMS Circuits with Bipolar Logarithmic Converter,” United States Patent 3,681,618, August 1, 1972.
3. Blackmer, “Multiplier Circuits,” United States Patent 3,714,462, January 30, 1973.
4. “Solid State Music Dual Linear-Antilog Voltage Controlled Amplifier,” SSM2000 Data Sheet, September, 1976.
5. Frey, “An Integrated Generalized Voltage Controlled Building Block”, *J. Audio Eng Soc.*, Preprint No. 2403, November 1986.
6. Curtis Electro Music Website, http://curtiselectromusic.com/Doug_Curtis.html.
7. Baskind, Rubens et al, “The Design and Integration of A High Performance Voltage Controlled Attenuator,” *J. Audio Eng Soc.*, Preprint No. 1555, November 1979.
8. Welland, “Compensation for VCA Op-Amp Errors,” United States Patent 4,434,380, February 28,1984
9. Analog Devices, “SSM2018T Trimless Voltage Controlled Amplifiers,” SSM2018T Datasheet, July 2002.
10. Blackmer, “Multiplier Circuits,” United States Patent 3,714,462, January 30, 1973.
11. Frey, “An Integrated Generalized Voltage Controlled Building Block,” *J. Audio Eng Soc.*, Preprint No. 2403, November 1986.
12. Frey, “Voltage-controlled element,” United States Patent 4,471,320, September 11, 1984.
13. Frey, “Monolithic Voltage-Controlled Element,” United States Patent 4,560,947, December 24, 1985.
14. Rubens, and Baskind, “Voltage Controlled Attenuator,” United States Patent 4,155,047, May 15, 1979.
15. “Solid State Music-Dual Linear-Antilog Voltage Controlled Amplifier,” SSM2000 Data Sheet, September 1976.
16. IEEE Std 152-1991-IEEE Standard for Audio Program Level Measurement, June 22, 1992.
17. Blackmer, “RMS Circuits with Bipolar Logarithmic Converter,” United States patent 3,681,618, August 1, 1972.
18. Chapel and Gurol, “Thermally Isolated Monolithic Semiconductor Die,” United States Patent 4,346,291, August 24, 1982.
19. Analog Devices, AD636 Datasheet.
20. Adams, “Networks for the Log Domain,” United States Patent 4,430,626, February 7, 1984.
21. THAT Corporation, “Adaptive Attack and Release Rates Using THAT Corporation RMS Detectors,” Design Note DN-114, 2000.
22. Jensen, “JE-990 Discrete Operational Amplifier,” *J. Audio Eng Soc.*, January/February 1980.
23. Jung, *IC Op-Amp Cookbook*, No. 20969, Howard W. Sams, 1974.
24. www.jensentransformers.com, May 2008.
25. Hebert and Thomas, “The 48 Volt Phantom Menace,” Audio Engineering Society Convention Paper 5335, May 2001.
26. Demrow, “Evolution From Operational Amplifier To Data Amplifier,” Analog Devices, 1968.
27. Wurcer, “A Programmable Instrumentation Amplifier for 12b Resolution Systems,” *International Solid State Circuits Conference*, February 1982. See also Analog Devices, Analog Dialog, Vol. 30 No. 2, 1996.
28. Harrison Systems, PC1041 Microphone Preamp Schematic, 1978.
29. Cohen, “Double Balanced Microphone Amplifier,” Audio Engineering Society Preprint 2106, September 1984.
30. Solid State Music, SSM2011 Datasheet, 1982.
31. Bowers, “An Ultra-Low-Noise Monolithic Microphone Preamplifier,” Audio Engineering Society Preprint 2495, 1987.
32. Whitlock, “A New Balanced Audio Input Circuit For Maximum Common-Mode Rejection In Real-World Environments,” www.jensen-transformers.com. 1996.
33. Hay, “Differential Technology In Recording Consoles And The Impact Of Transformerless Circuitry On Grounding Technique,” Audio Engineering Society Preprint 1723, October 1980.
34. Whitlock, “Differential Line Receiver With Common-Mode Bootstrapping,” US patent 5,568,561, October 22, 1996.
35. Strahm, “Balanced Line Output Circuit,” United States Patent 4,979,218, December, 18, 1990.

Chapter 13

Heatsinks and Relays

by Glen Ballou and Henry Villaume

13.1 Heatsinks 367
  13.1.1 Thermal Management of Today's Audio Systems 367
    13.1.1.1 Convection 367
    13.1.1.2 Conduction 367
    13.1.1.3 Radiation 368
    13.1.1.4 Summary 368
  13.1.2 New Technologies to Make Things Fit More Easily. 368
  13.1.3 How Heatsinks Work 369
    13.1.3.1 Thermal Resistance 370
    13.1.3.2 Heatsink Materials and Design 370
13.2 Relays 373
  13.2.1 Glossary of Terms 374
  13.2.2 Contact Characteristics6,7,8 377
  13.2.3 Relay Loads[8] 378
    13.2.3.1 The Effects of Various Loads 379
  13.2.4 Electromechanical Relays 379
    13.2.4.1 Dc Relays 380
    13.2.4.2 Ac Relays 380
  13.2.5 Reed Relays[6,7,8,11] 381
    13.2.5.1 Contact Resistance and Dynamics 381
    13.2.5.2 Magnetic Interaction 381
    13.2.5.3 Environmental Temperature Effects 382
    13.2.5.4 Dry Reed Relays 382
    13.2.5.5 Mercury-Wetted Contact Relays 384
    13.2.5.6 RF Relays 384
    13.2.5.7 Dry Reed Switches 387
  13.2.6 Solid-State Relays9 388
    13.2.6.1 Advantages 388
    13.2.6.2 Disadvantages and Protection4 388
    13.2.6.3 High-Peak-Transient-Voltage Protection 389
    13.2.6.4 Low Load Current Protection 390
    13.2.6.5 Optically Coupled Solid-State Relays 390
    13.2.6.6 Transformer-Coupled Solid-State Relays 390
    13.2.6.7 Direct-Coupled Solid-State Relays 390
    13.2.6.8 Solid-State Time-Delay Relays10 391
References 392

This page intentionally left blank

13.1 Heatsinks

13.1.1 Thermal Management of Today's Audio Systems

By Henry Villaume, Villaume Associates, LLC

Today's audio systems, like all electronic systems are being powered by smaller devices, packaged in smaller systems that are generating more heat. We need to increase our level of understanding on all of the latest techniques for the management of this added heat in as effective a means as possible. Let's first start with the understanding of the three methods of heat transfer—specifically, convection, conduction, and radiation as all three methods of heat transfer contribute to the complete thermal management provided by the heatsinks installed in an audio system.

13.1.1.1 Convection

Convection is the transfer of heat from a solid surface to the surrounding gas, which is always air in the case of a typical audio system. This method of heat transfer drives the amount of required fin surface area that is accessible to the surrounding air so that it may heat up the surrounding air, allow it to move away, and make room for the process to repeat itself. This process can be greatly accelerated with the use of a fan to provide more energy to the moving of the air than just the natural buoyant force of the heated air.

Natural convection is when there is no external fan and the heat transfer occurs with very low air flow rates, typically as low as 35 linear feet per minute (lfm) for obstructed natural convection to 75 lfm for optimum unobstructed vertical natural convection. Natural convection is never zero air flow rate because without air movement there would be no heat transfer. Think of the closed cell plastic foam insulation. It works as an insulator because the closed cell prevents the air from moving away.

Forced convection is when a system fan imparts a velocity to the air surrounding the heatsink fins. The fan may be physically attached to the convective fin surface area of the heatsink to increase the air velocity over the fin surfaces. There is impingement flow—fan blows down from on top of the fins—and through flow—fan blows from the side across the fin set.

Forced convection thermal systems are most generally significantly smaller (50% or more) than their natural convection equivalents. The penalties for the smaller size are the added power to operate the fan, an added failure mechanism, the added cost, and the noise from the fan. Fan noise is probably the most important consideration when applying them in audio systems.

13.1.1.2 Conduction

Conduction is the transfer of heat from one solid to the next adjacent solid. The amount and thermal gradient of heat transfer are dependent on the surface finishes—flatness and roughness—and the interfacial pressure generated by the attachment system. This mechanically generated force is accomplished by screws, springs, snap assemblies, etc. The thermal effectiveness of a conductive interface is measured by the resultant temperature gradient in °C. This may be calculated from the interface thermal resistance at the mounted pressure times the watts of energy moving across the joint divided by the cross-sectional area. These temperature gradients are most significant for high wattage components in small packages—divisors less than 1.0 are, in actual effect, multipliers. Good thermal solutions have attachment systems that generate pressures of 25–50 psi.

Table 13-1 compares the thermal performance of most of the common interface material groups with a dry joint—this makes amply clear why it is never acceptable to specify or default through design inaction to a dry joint.

Table 13-1. Thermal Performance of Common Interface Materials

Interface Material Group	Thermal Performance Range in °C in^2/W	Comments
Dry Mating Surfaces	3.0-12.0	Too much uncertainty to use. Too big a thermal gradient
Gap Fillers	0.4-4.0	Minimize thickness required Spring mechanical load
Electrically Insulating	0.2- 1.5	Maximize mechanical loads
High Performance Pads	0.09-0.35	Minimize thickness Maximize mechanical loads
Phase Change Pads	0.02-0.14	Must follow application method Spring mechanical load
Low Performance Grease	0.04-0.16	Screen apply Spring mechanical load
High Performance Grease	0.009-0.04	Must Screen apply Spring mechanical load Best at high loads(> 50 psi)

The primary heat transfer driving force is the temperature difference between $T_{maxcase}$ and $T_{maxambient}$ modified by the conductive delta T losses in the interface and any extraordinary hot spot offsets and spreading losses. If heatsinks are mounted over spread hot spots these last conductive losses are not sufficiently large to consider. They really only become significant when considering very unusual arrangements—high-watt density loads such as typical LEDs.

13.1.1.3 Radiation

Radiation is the third and least important method of heat transfer for audio system heatsinks. Radiation has a maximum 20–25% impact in natural convection applications with a negligible impact after 200 lfm applications. Radiation is a function of the fourth power of the absolute temperature difference between the hot side and surrounding cooler surfaces that look at each other and their respective emissivities. In the real world in which we live, these are not significant enough to suggest a lot of effort to understand and optimize.

Aluminum extruded heatsinks were typically made black in an anodizing process at a significant cost to get an emissivity of 0.95 (dimensionless). The typical aluminum surface forms an oxide film in less than a second after machining with an emissivity of about 0.30–0.40. Nominally, almost half the benefit will come free so that our advice is “Leave the radiation effects alone.” What you get beneficially you were going to largely get anyway, free from Mother Nature.

In review, heatsinks use all three methods of heat transfer to produce the desired effect of cooling the typical electronic component in the typical audio system.

13.1.1.4 Summary

Convection is usually the most significant method, and it depends on having sufficient fin surface area in direct contact with the surrounding air and design features to minimize the insulating effects of boundary films. Aerodynamic shapes and adequate open fin spacing that allows free air movement are critical design issues.

Conduction is the first step in the heat transfer chain in that conduction transfers the heat from the device into heatsink, then through the heatsink to the fin surface where convection takes over. Some heatsinks need conduction enhancements such as heat pipes to keep the conduction temperature gradients to a value that is low enough to allow the convection to complete the heat transfer without exceeding the application temperature limits.

Radiation is a secondary level effect that is always present, marginally significant in natural convection, but not economical to control.

13.1.2 New Technologies to Make Things Fit More Easily

The range of technologies, materials, and fabrication processes available to the thermal designer today is quite impressive. The primary goal when employing these advanced technologies, materials, and fabrication processes is to increase the effective density of the of the resultant heat transfer system. Technically, we are increasing the volumetric efficiency of the thermal solution proposed for a given application. In “man speak” the required heatsink gets much smaller in size and therefore fits more easily into the ever-shrinking product envelope. A smaller heatsink has a decreased conductive thermal spreading resistance and therefore a smaller conductive temperature gradient. In this section we are going to assume that we have a convective solution defined for a baseline heatsink. The baseline heatsink is fabricated from an extruded aluminum alloy (6063-T5). The following paragraphs will describe a technology, material or fabrication process and give a volume ratio or range of volume ratios that can be applied to the existing solution to quickly see the benefit of applying this technology, material, or fabrication process to the audio application at hand. Ratios that are less than 1.00 are indicating a reduction in heatsink volume.

Thermal solution problem solving is an iterative process balancing the application boundary specifications against the affordable technologies/materials/fabrication processes until a system compromise solution is defined.[1] For example; marketing has directed that only a natural convection solution is acceptable but the heatsink is too big. One solution might require the $T_{maxambient}$ be reduced by 5°C and the heatsink be fabricated from copper, C110 soldered together. This could reduce the size of the heatsink by 25–35%. The penalties would be the weight would increase between to and three times and the unit cost of the heatsink increase by three to four times. There are software systems[2] that specialize in defining these trade-offs rapidly, allowing a real-time compromise to be made, even during the design review meeting with marketing.

Table 13-2 summarizes the thermal solution benefits possible with the proper application of new technologies, materials and fabrication processes.

Extruded heatsinks have fin thicknesses that are much greater, thicker than required thermally. They are thicker to accommodate the strength requirements of the die, which is close to the melting point of aluminum during the extrusion process.

Bonded fin and folded fin heatsink designs use sheet stock for the fins so that they may be optimally sized as required to carry the thermal load without regard to the mechanical requirements of the extrusion process. These heatsinks can, therefore, without compromising the required open fin spacing, have a greater number of fins and be convectively much more volumetrically effective. These sheet metal fins are attached to the heatsink bases with either thermal epoxy adhesives or solders. Since this joint only represents ~3 % of the total thermal resistance of the heatsink, the adhesive choice is never critical.

Air flow management is the most critical parameter to control in optimizing the convective heat transfer for any thermal solution. Baffles, shrouds, and fan sizing are all very critical in making the most of the convective portion of the heat transfer thermal solution. Some months ago we were confronted with an audio amplifier that had two rows of very hot components. With two facing extrusions that formed a box shape, we mounted a fan at the end and blew the air down the chute with great success. The air flow was fully contained and no leakage occurred. And so the audio cooling tube was born.

There should always be a space—0.5–0.8 inches along the axis of the fan—between the fan outlet and the fin set that is fully shrouded to force the air to pass over the convective fin surfaces. This is called the *plenum*. Its function is to allow the upstream pressure generated by the fan to reach an equilibrium and thereby equalize the air flow through each fin opening.

Audio systems that require fans need to be carefully designed to have an air flow path that is well defined so that the fan may be operated at a minimal speed. This results in the fan generating a minimum of noise. High-velocity fans are noisy. Noise abatement is very expensive and seldom truly satisfactory, therefore, the best solution is to minimize the fan generated noise.

13.1.3 How Heatsinks Work

by Glen Ballou

Heatsinks are used to remove heat from a device, often the semiconductor junction. To remove heat, there must be a temperature differential (ΔT) between the junction and the air. For this reason, heat removal is always after the fact. Unfortunately, there is also resistance to heat transfer between the junction and its case, any insulating material, the heatsink, and the air, Fig. 13-1.

Table 13-2. Thermal Solutions with New Technology, Materials, and Fabrication Processes

Technology/ Material/ Fabrication	Title	Volumetric Ratio Range	Cost Range	Comments
M	Copper C110	0.8	3.5 ×	Volumetric ratios are even lower for conduction=limited applications Weight almost triples (3 ×)
MF	Molded Plastic Conductive Dielectric Elastomeric	0.97 (<200 LFM) 1.03 (>200 LFM) 1.07 (>500 LFM)	0.5–0.7 After tooling if not a standard	Saves weight and finishing[3] Hybrids—molded base metal fins
TM	Base-Mounted Heat Pipes[3]	1 Heat Pipe 0.79 2 Heat Pipe 0.73	1.5–2.0	Aluminum base
TM	Base-Mounted Heat Pipes[3]	1 Heat Pipe 0.71 2 Heat Pipe 0.66	4.0–4.5	Copper base
TM	Base-Mounted Vapor Chamber[3]	0.69 Al 0.58 Cu	2.5 Al 4.8 Cu	Achieves optimum spreading
M	Graphite	0.72	6–8 ×	Relatively fragile 35% reduction in weight
TM	Solid-State Heat Pipes (TPG)[4]	0.75	3.6 Al 5.4 Cu	Eliminates burnout as a failure mode
FM	Bonded Fin and Folded Fin	0.90 Al 0.76 Cu	2 × 3.6 ×	More convective fin surface per unit volume fin shapes break up boundary film layers for performance gains

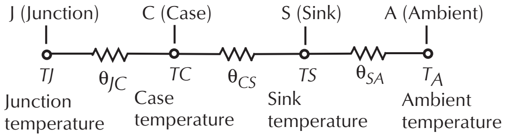
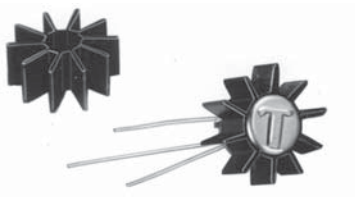
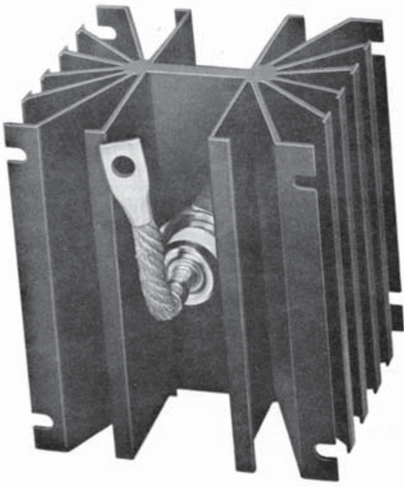


Figure 13-1. Series thermal resistance/temperature circuit.



A. Small heat sinks used with diodes and transistors. Their diameter is less than a dime.



B. Large heat sink for use with heavy current rectifiers. The stud of the rectifier is screwed into the center fin of the sink.

Figure 13-2. Conduction type heatsinks used for cooling diodes and transistors. Courtesy Wakefield Engineering Co.

13.1.3.1 Thermal Resistance

The total thermal resistance between the junction and the air is the sum of the individual thermal resistances

$$\Sigma\theta = \theta_{JC} + \theta_{CI} + \theta_{IS} + \theta_{SA} \tag{13-1}$$

where,
I is the thermal resistance in degrees Celsius per watt (°C/W),
JC is the junction to case,
CI is the case to insulator,
IS is the insulator to heatsink,
SA is the heatsink to air.

The temperature at the junction can be determined from the ambient temperature, the thermal resistance between the air and the junction, and the power dissipated at the junction.

$$T_J = T_A + \theta_{JA} P_D \tag{13-2}$$

where,
T_A is the temperature of the air,
θ_{JA} is the thermal resistance from the air to the junction,
P_D is the power dissipated.

If the junction temperature was known, then the power dissipated at the junction can be determined

$$P_D = \frac{\Delta T}{\Sigma\theta} \tag{13-3}$$

where,
ΔT is $T_J - T_A$.

13.1.3.2 Heatsink Materials and Design

Heatsinks are generally made from extruded aluminum or copper and are painted black, except for the areas in which the heat-producing device is mounted. The size of heatsinks will vary with the amount of heat to be radiated and the ambient temperature and the maximum average forward current through the element. Several different types of heatsinks are pictured in Fig. 13-2.

The rate of heat flow from an object is

$$Q = \frac{KA\Delta T}{L} \tag{13-4}$$

where,
Q is the rate of heat flow,
K is the thermal conductivity of material,
A is the cross-sectional area,
ΔT is the temperature difference,
L is the length of heat flow.

For best conduction of heat, the material should have a high thermal conductivity and have a large cross-sectional area. The ambient or material temperature should be maintained as low as possible, and the thermal path should be short.

The heat may also be transferred by convection and radiation. When a surface is hotter than the air about it, the density of the air is reduced and rises, taking heat with it. The amount of heat (energy) radiated by a body is dependent on its surface area, temperature, and emissivity. For best results, the heatsink should:

- Have maximum surface area/volume (hence the use of vertical fins).

- Be made of a high thermal conductivity material.
- Have material of high emissivity (painted aluminum or copper).
- Have proper ventilation and location (should be below, not above, other heat radiators).
- Be placed so that the lowest power device is below the higher power devices, and all devices should be as low as possible on the heatsink.

The overall effectiveness of a heatsink is dependent to a great extent on the intimacy of the contact between the device to be cooled and the surface of the heatsink. Intimacy between these two is a function of the degree of conformity between the two surfaces and the amount of pressure that holds them together. The application of a silicone oil to the two surfaces will help to minimize air gaps between the surfaces, improving conduction. The use of a mica washer between the base of the device to be cooled and the heatsink will add as much as 0.5°C/W to the thermal resistance of the combination. Therefore, it is recommended that (whenever possible) an insulating washer be used to insulate the entire heatsink from the chassis to which it is to be mounted. This permits the solid-state device to be mounted directly to the surface of the heatsink (without the mica washer). In this way, the thermal resistance of the mica washer is avoided.

Today high thermal conductive/high electrical insulation materials are available to electrically insulate the transistor case from the heatsink. They come in the form of silicon rubber insulators, hard-coat-anodized finish aluminum wafers, and wafers with a high beryllium content.

A typical heatsink is shown in Fig. 13-3. This sink has 165 in^2 of radiating surface. The graph in Fig. 13-4 shows the thermal characteristics of a heatsink with a transistor mounted directly on its surface. A silicone oil is used to increase the heat transfer. This graph was made with the heatsink fins in a vertical plane, with air flowing from convection only. Fig. 13-5 shows the effect of thermal resistance with forced air blown along the length of the fin.

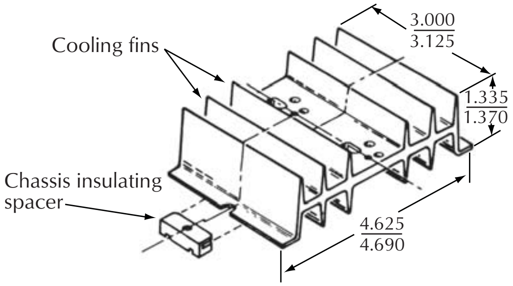


Figure 13-3. Typical heatsink for mounting two transistors. Courtesy Delco Electronics Corp.

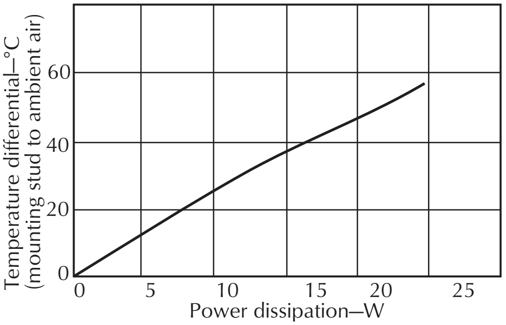


Figure 13-4. Thermal characteristics for the heatsink shown in Fig. 13-3, with forced air cooling. Courtesy of Delco Electronics Corp.

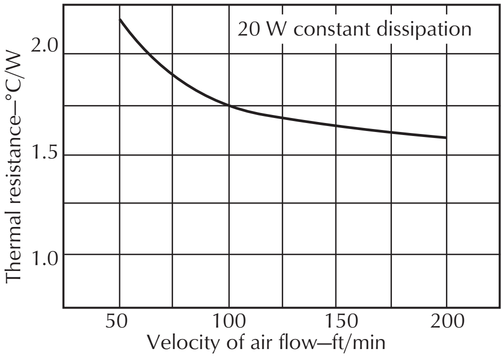


Figure 13-5. Thermal characteristics for the heatsink shown in Fig. 13-3, with forced air cooling. Courtesy of Delco Electronics Corp.

A transistor mounting kit is shown in Fig. 13-6. Several different types of silicon fluids are available to improve heat transfer from the device to the heatsink. The fluid is applied between the base of the transistor and the surface of the heatsink or, if the transistor is insulated from the heatsink, between the base and the mica washer and the mica washer and the heatsink. For diodes pressed into a heatsink, the silicone fluid is applied to the surface of the diode case before pressing it into the heatsink. The purpose of the silicone fluid is to provide good heat transfer by eliminating air gaps.

Thermally conductive adhesives can also be used. These adhesives offer a high heat transfer, low shrinkage, and a coefficient of thermal expansion comparable to copper and aluminum.

The thermal capacity of a cooling fin or heatsink must be large compared to the thermal capacity of the

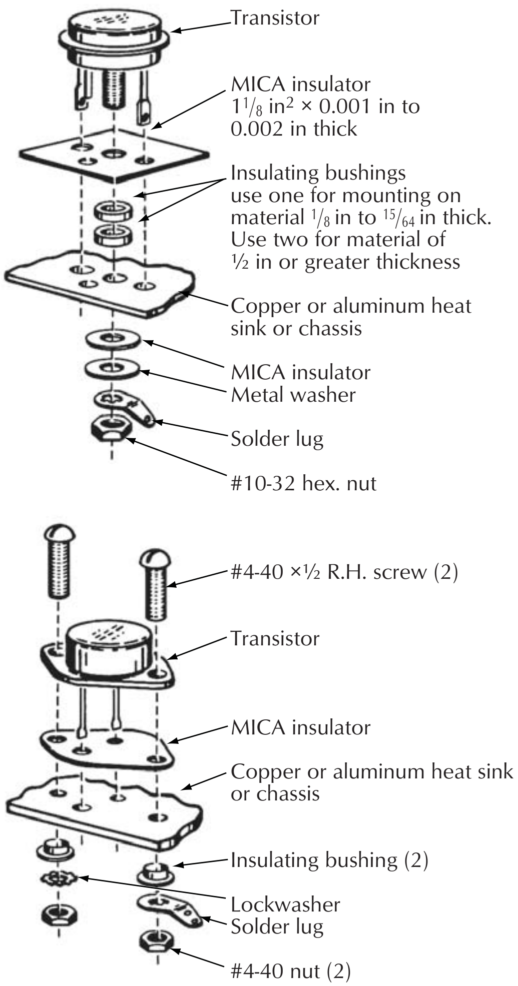

Transistor
MICA insulator
1 1/8 in^2 × 0.001 in to
0.002 in thick
Insulating bushings
use one for mounting on
material 1/8 in to 15/64 in thick.
Use two for material of
½ in or greater thickness
Copper or aluminum heat
sink or chassis
MICA insulator
Metal washer
Solder lug
#10-32 hex. nut
#4-40 ×½ R.H. screw (2)
Transistor
MICA insulator
Copper or aluminum heat sink
or chassis
Insulating bushing (2)
Lockwasher
Solder lug
#4-40 nut (2)

Figure 13-6. Transistor/heatsink mounting kits. Courtesy Delco Electronics Corp.

Table 13-3. Specific Thermal Resistance *p* of Interface Materials, °C in/W

Material	ρ
Still air	1200
Mylar film	236
Silicone grease	204
Mica	66
Wakefield Type 120 Compound	56
Wakefield Delta Bond 152	47
Anodize	5.6
Aluminum	0.19
Copper	0.10

Courtesy of Wakefield Engineering, Inc.

device and have good thermal conductivity across its entire area. The specific thermal resistance ρ of interface materials used for heatsinks and insulating devices is shown in Table 13-3.

The thermal resistance θ for these materials can be determined by the equation

$$\theta = \frac{\rho t}{A} \tag{13-5}$$

where,

ρ is the specific thermal resistance,

t is the material thickness in inches,

A is the area in square inches.

For instance, a square copper plate 4 inches per side and ⅛ inch thick would have a θ of 0.00078°C/W, while a mica insulator 0.003 inch thick with a diameter of 1 inch would have a θ of 0.25°C/W. If the semiconductor dissipates 100 W, the temperature drop across the copper plate would be 0.07°C (0.13°F) and across the mica washer it would be 25°C (45°F). In transistor replacement in older equipment, it would be best to replace the mica insulator with a new type of insulator.

In the selection of a heatsink material, the thermal conductivity of the material must be considered. This determines the thickness required to eliminate thermal gradients and the resultant reduction in emissivity. An aluminum fin must be twice as thick as a comparable copper fin, and steel must be eight times as thick as copper.

Except for the smallest low-current solid-state devices, most devices must use a heatsink, either built in or external.

Space for heatsinks is generally limited, so the minimum surface area permissible may be approximately calculated for a flat aluminum heat plate by

$$A = 133\frac{W}{\Delta T}\text{ in}^2 \tag{13-6}$$

where,

W is the power dissipated by the device,

ΔT is the temperature differences between the ambient and case temperature in °C.

The approximate wattage dissipated by the device can be calculated from the load current and the voltage drop across it

$$W = I_L V_D \tag{13-7}$$

where,

I_L is the load current,

V_D is the voltage drop across the device.

For a triac, V_D is about 1.5 V; for SCRs, about 0.7 V. For transistors it could be from 0.7 V to more than 100 V.

The following is an example of how to determine the minimum surface area required for a flat aluminum heatsink to keep the case temperature of 75°C (167°F) for a triac while delivering a load current of 15 A, at 25°C (77°F) ambient and a voltage drop across the triac of 1.5 V

$$\begin{aligned}\Delta T &= T_{case} - T_{ambient}\\ &= 75°C - 25°C\\ &= 50°C\end{aligned}$$

Using Eq. 13-7

$$\begin{aligned}W &= V_D I_L\\ &= 1.5 \times 15\\ &= 22.5\ \text{W}\end{aligned}$$

Using Eq. 13-6

$$\begin{aligned}A &= 133\frac{W}{\Delta T}\ \text{in}^2\\ &= 133 \times \frac{22.5}{50}\\ &= 59.85\ \text{in}^2\end{aligned}$$

It is important that the case temperature, T_{case}, does not exceed the maximum allowed for a given load current, I_L (see typical derating curves in Fig. 13-70).

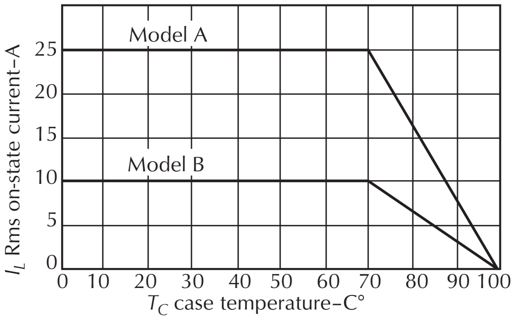


Figure 13-7. A typical derating curve for solid state devices.

Eq. 13-6 gives the surface area needed for a vertically mounted heatsink. With free air convection, a vertically mounted heatsink, Fig. 13-8, has a thermal resistance approximately 30% lower than with horizontal mounting.

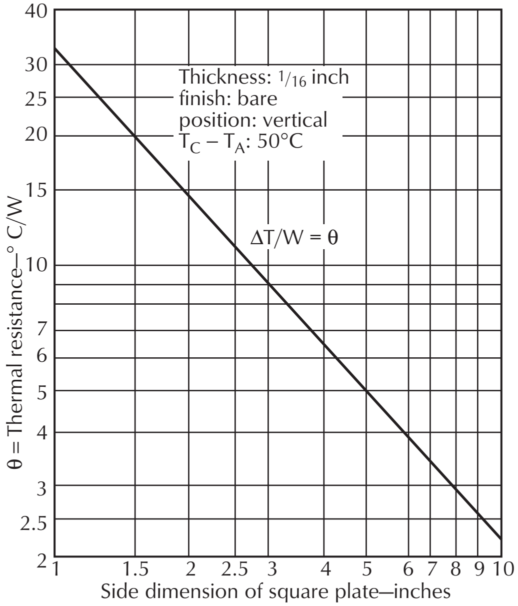


Figure 13-8. Thermal resistance for a vertically mounted 1⁄16 inch aluminum plate of various dimensions.

In restricted areas, forced-convection cooling may be necessary to reduce the effective thermal resistance of the heatsink. When forced air cooling is used to cool the component, the cubic feet per minute (cfm or ft^3/min) required is determined by

$$\begin{aligned}cmf &= \frac{Btu/h}{60} \times 0.02\ \text{temperature rise}\\ &= \frac{1.76Q}{\Delta TK}\end{aligned} \tag{13-8}$$

where,

1 W is 3.4 Btu,

temperature rise is in °C,

Q is the heat dissipated in watts,

ΔT is the heatsink mounting temperature minus the ambient temperature,

K is the coupling efficiency (0.2 for wide spaced fins, 0.6 for close spaced fins).

13.2 Relays

A *relay* is an electrically operated switch connected to or actuated by a remote circuit. The relay causes a second circuit or group of circuits to operate. The relay may control many different types of circuits connected to it. These circuits may consist of motors, bells, lights, audio circuits, power supplies, and so on, or the relay may be used to switch a number of other circuits at the same time or in sequence from one input.

Relays may be electromechanical or solid state. Both have advantages and disadvantages. Only a few years ago relays were big and cumbersome and required either an octal-type socket or were externally wired. Today relays are very compact and come in many layouts. A few are given below.

Solder Connectors. Connectors vary in size and spacing, depending on the current carrying capacity.

Octal Sockets. Plug into standard 8 pin and 11 pin sockets.

Rectangular Sockets. Plug into a 10 pin, 11 pin, 12 pin, 14 pin, 16 pin, 22 pin, or 28 pin socket.

DIP Relays. Designed to mount directly on a printed circuit board on a 0.1 inch spacing. Sockets can be 8 pin or 16 pin.

SIP 4 Pin Relay. Plug into a SIP socket or mount on a printed circuit board on a 0.2 inch in-line spacing.

13.2.1 Glossary of Terms

This glossary was compiled from NARM Standard RS-436, MIL STD 202, and MIL STD R5757.

Actuate Time. The time measured from coil energization to the stable contact closure (Form A) or stable contact opening (Form B) of the contact under test. (See also Operate Time.)

Ampere Turns (AT). The product of the number of turns in an electromagnetic coil winding and the current in amperes passing through the winding.

Bandwidth. The frequency at which the RF power insertion loss of a relay is 50%, or –3 dB.

Bias, Magnetic. A steady magnetic field applied to the magnetic circuit of a switch to aid or impede its operation in relation to the coil's magnetic field.

Bounce, Contact. Intermittent and undesired opening of closed contacts or closing of opened contacts usually occurring during operate or release transition.

Breakdown Voltage. The maximum voltage that can be applied across the open switch contacts before electrical breakdown occurs. In reed relays it is primarily dependent on the gap between the reed switch contacts and the type of gas fill used. High AT switches within a given switch family have larger gaps and higher breakdown voltage. It is also affected by the shape of the contacts, since pitting or whiskering of the contact surfaces can develop regions of high electric field gradient that promote electron emission and avalanche breakdown. Since such pitting can be asymmetric, breakdown voltage tests should be performed with forward and reverse polarity. When testing bare switches, ambient light can affect the point of avalanche and should be controlled or eliminated for consistent testing. Breakdown voltage measurements can be used to detect reed switch capsule damage. See Paschen Test.

Carry Current. The maximum continuous current that can be carried by a closed relay without exceeding its rating.

Coaxial Shield. Copper alloy material that is terminated to two pins of a reed relay within the relay on each side of the switch. Used to simulate the outer conductor of a coaxial cable for high-frequency transmission.

Coil. An assembly consisting of one or more turns of wire around a common form. In reed relays, current applied to this winding generates a magnetic field that operates the reed switch.

Coil AT. The coil *ampere turns* (AT) is the product of the current flowing through the coil (and therefore directly related to coil power) and the number of turns. The coil AT exceeds the switch AT by an appropriate design margin to ensure reliable switch closure and adequate switch overdrive. Sometimes abbreviated as NI, where N is the number of turns and I is the coil current.

Coil Power. The product, in watts, of the relay's nominal voltage and current drawn at that voltage.

Cold Switching. A circuit design that ensures the relay contacts are fully closed before the switched load is applied. It must take into account bounce, operate and release time. If technically feasible, cold switching is the best method for maximizing contact life at higher loads.

Contact. The ferromagnetic blades of a switch often plated with rhodium, ruthenium, or tungsten material.

Contact Resistance, Dynamic. Variation in contact resistance during the period in which contacts are in motion after closing.

Contact Resistance, Static. The dc resistance of closed contacts as measured at their associated contact terminals. Measurement is made after stable contact closure is achieved.

Crosstalk (Crosstalk Coupling). When applied to multichannel relays, the ratio, expressed in dB, of the signal power being emitted from a relay output contact to the power being applied to an adjacent input channel at a specified frequency.

Duty Cycle. A ratio of energized to de-energized time.

Electrostatic Shield. Copper alloy material terminated to one pin within the reed relay. Used to minimize coupling and electrostatic noise between the coil and contacts.

Form-A. Contact configuration that has one single pole–single throw normally open (SPST n.o.) contact.

Form-B. Contact configuration that has one single pole–single throw normally closed (SPST n.c.) contact.

Form-C. Contact configuration that has one single pole–double throw (SPDT) contact. (One common point connected to one normally open and one normally closed contact.) Sometimes referred to as a transfer contact.

Hard Failure. Permanent failure of the contact being tested.

Hermetic Seal. An enclosure that is sealed by fusion to ensure a low rate of gas leakage. In a reed switch, a glass-to-metal seal is employed.

Hot Switching. A circuit design that applies the switched load to the switch contacts at the time of opening and closure.

Hysteresis. When applied to reed relays, the difference between the electrical power required to initially close the relay and the power required to just maintain it in a closed state. (Usually expressed in terms of the relay's pull-in voltage and drop-out voltage.) Some degree of hysteresis is desirable to prevent chatter and is also an indicator of adequate switch contact force.

Impedance (Z). The combined dc resistance and ac reactance of a relay, at a specified frequency and if found with the equation

$$Z = R + jX \tag{13-9}$$

where,

R is the dc resistance,

X is $2\pi fL - \frac{1}{2\pi fC}$,

f is the frequency.

Because of the small residual capacitance across the open contacts of a reed relay, the impedance decreases at higher frequencies, resulting in lower isolation at higher frequencies. Conversely, increasing inductive reactance at higher frequencies causes the impedance of a closed relay to rise, increasing the insertion loss at higher frequencies.

Impedance Discontinuity. A deviation from the nominal RF impedance of 50 Ω at a point inside a reed relay. Impedance discontinuities cause signal absorption and reflectance problems resulting in higher signal losses. They are minimized by designing the relay to have ideal transmission line characteristics.

Insertion Loss. The ratio of the power delivered from an ac source to a load via a relay with closed contacts, compared to the power delivered directly, at a specified frequency, and is found with the equation

$$Insertion\ Loss = -20\log\frac{V_t}{V_i} \tag{13-10}$$

where,

V_t is the transmitted voltage,

V_i is the incident voltage.

Insertion loss, isolation and return loss are often expressed with the sign reversed; for example, the frequency at which 50% power loss occurs may be quoted as the –3 dB point. Since relays are passive and always produce net losses, this does not normally cause confusion.

Inrush Current. Generally, the current waveform immediately after a load is connected to a source. Inrush current can form a surge flowing through a relay that is switching a low-impedance source load that is typically a highly reactive circuit or one with a nonlinear load characteristic such as a tungsten lamp load. Such abusive load surges are sometimes encountered when relays are inadvertently connected to test loads containing undischarged capacitors or to long transmission lines with appreciable amounts of stored capacitive energy. Excessive inrush currents can cause switch contact welding or premature contact failure.

Insulation Resistance. The dc resistance between two specified test points.

Isolation. The ratio of the power delivered from a source to a load via a relay with open contacts, compared to the power delivered directly, at a specified frequency. If V_i is the incident voltage and V_t is the transmitted voltage, then isolation can be expressed in decibel format as

$$Isolation = -20\log\frac{V_t}{V_i} \tag{13-11}$$

where,
V_t is the transmitted voltage,
V_i is the incident voltage.

Latching Relay. A bistable relay, typically with two coils, that requires a voltage pulse to change state. When pulse is removed from the coil, the relay stays in the state in which it was last set.

Life Expectancy. The average number of cycles that a relay will achieve under specified load conditions before the contacts fail due to sticking, missing or excessive contact resistance. Expressed as mean cycles before failure (MCBF).

Low Thermal Emf Relay. A relay designed specifically for switching low-voltage level signals such as thermocouples. These types of relays use a thermally compensating ceramic chip to minimize the thermal offset voltage generated by the relay.

Magnetic Interaction. The tendency of a relay to be influenced by the magnetic field from an adjacent energized relay. This influence can result in depression or elevation of the pull-in and dropout voltage of the affected relay, possibly causing them to fall outside their specification. Magnetic interaction can be minimized by alternating the polarity of adjacent relay coils, by magnetic shielding, or by placing two relays at right angles to each other.

Magnetic Shield. A ferromagnetic material used to minimize magnetic coupling between a relay and external magnetic fields.

Mercury Wetted Contact. A form of reed switch in which the reeds and contacts are wetted by a film of mercury obtained by a capillary action from a mercury pool encapsulated within the reed switch. The switch in this type of relay must be mounted vertically to ensure proper operation.

Missing (Contacts). A reed switch failure mechanism, whereby an open contact fails to close by a specified time after relay energization.

Nominal Voltage. The normal operating voltage of the relay.

Operate Time. The time value measured from the energization of the coil to the first contact closure, Form A, or the first contact open, Form B.

Operate Voltage. The coil voltage measured at which a contact changes state from its unenergized state.

Overdrive. The fraction or percentage by which the voltage applied to the coil of a relay exceeds its pull-in voltage. An overdrive of at least 25% ensures adequate closed contact force and well-controlled bounce times, which result in optimum contact life. For instance, Coto Technology's relays are designed for a minimum of 36% overdrive so a relay with a nominal coil voltage of 5 V will pullin at no greater than 3.75 V.

When using reed relays, the overdrive applied to the relay should not drop below 25% under field conditions. Issues such as power supply droop and voltage drops across relay drivers can cause a nominally acceptable power supply voltage to drop to a level where adequate overdrive is not maintained.

Release Time. The time value measured from coil de-energization to the time of the contact opening, Form-A or first contact closure, Form-B.

Release Voltage. The coil voltage measured at which the contact returns to its de-energized state.

Return Loss. The ratio of the power reflected from a relay to that incident on the relay, at a specified frequency and can be found with the equation

$$Return\ loss = -20\log\frac{V_r}{V_i} \tag{13-12}$$

where,
V_r is the reflected voltage,
V_i is the incident voltage.

Signal Rise Time. The rise time of a relay is the time required for its output signal to rise from 10– 90% of its final value, when the input is changed abruptly by a step function signal.

Shield, Coaxial. A conductive metallic sheath surrounding a reed relay's reed switch, appropriately connected to external pins by multiple internal connections, and designed to preserve a 50 Ω impedance environment within the relay. Used in relays designed for high-frequency service to minimize impedance discontinuities.

Shield, Electrostatic. A conductive metallic sheath surrounding a reed relay's reed switch, connected to at least one external relay pin, and designed to minimize capacitive coupling between the switch and other relay components, thus reducing high-frequency noise pickup. It is similar to a coaxial shield, but not designed to maintain a 50 Ω RF impedance environment.

Shield, Magnetic. An optional plate or shell constructed of magnetically permeable material such as nickel-iron or mu-metal, fitted external to the relay's coil. Its function is to reduce the effects of magnetic interaction between adjacent relays and to improve the efficiency of the relay coil. A magnetic shell also reduces the influence of external magnetic fields, which is useful in security applications. Magnetic shields can be fitted externally or may be buried inside the relay housing.

Soft Failure. Intermittent self-recovering failure of a contact.

Sticking (Contacts). A switch failure mechanism, whereby a closed contact fails to open by a specified time after relay de-energization. Can be subclassified as hard or soft failures.

Switch AT. The ampere turns required to close a reed switch, pull-in AT, or just to maintain it closed, drop-out AT, and is specified with a specific type and design of coil. Switch AT depends on the length of the switch leads and increases when the reed switch leads are cropped. This must be taken into account when specifying a switch for a particular application.

Switching Current. The maximum current that can be hot-switched by a relay at a specified voltage without exceeding its rating.

Switching Voltage. The maximum voltage that can be hot-switched by a relay at a specified current without exceeding its rating. Generally lower than breakdown voltage, since it has to allow for any possible arcing at the time of contact breaking.

Transmission Line. In relay terms an interruptible waveguide consisting of two or more conductors, designed to have a well-controlled characteristic RF impedance and to efficiently transmit RF power from source to load with minimum losses, or to block RF energy with minimum leakage. Structures useful within RF relays include microstrips, coplanar waveguides, and coaxial transmission line elements.

VSWR (Voltage Standing Wave Ratio). The ratio of the maximum RF voltage in a relay to the minimum voltage at a specified frequency and calculated from

$$VSWR = (1+\rho)/(1-\rho) \qquad (13\text{-}13)$$

where,

ρ is the the voltage reflected back from a closed relay terminated at its output with a standard reference impedance, normally 50 Ω.

13.2.2 Contact Characteristics[6,7,8]

Contacts may switch either power or dry circuits. A power circuit always has current flowing, while a dry circuit has minimal or no current flowing, such as an audio circuit. A dry or low-level circuit typically is less than 100 mV or 1 mA.

The mechanical design of the contact springs is such that when the contacts are closed, they slide for a short distance over the surfaces of each other before coming to rest. This is called a *wiping contact*, and it ensures good electrical contact.

Contacts are made of silver, palladium, rhodium, or gold and may be smooth or bifurcated. Bifurcated contacts have better wiping and cleaning action than smooth contacts and, therefore, are used on dry circuits.

There are various combinations of contact springs making up the circuits that are operated by the action of the relay. Typical spring piles are shown in Fig. 13-9.

As contacts close, the initial resistance is relatively high, and any films, oxides, and so on further increase the contact resistance. Upon closing, current begins to flow across the rough surface of the contacts, heating and softening them until the entire contact is mating, which reduces the contact resistance to milliohms. When the current through the circuit is too low to heat and soften the contacts, gold contacts should be used since the contacts do not oxidize and, therefore, have low contact resistance. On the other hand, gold should not be used in power circuits where current is flowing.

The contact current specified is the maximum current, often the make-or-break current. For instance, the make current of a motor or capacitor may be 10–15 times as high as its steady-state operation. Silver cadmium oxide contacts are very common for this type of load. The contact voltage specified is the maximum voltage allowed during arcing during break. The break voltage of an inductor can be 50 times the steady-state voltage of the circuit.

To protect the relay contacts from high transient voltages, arc suppression should be used. For dc loads, this may be in the form of a reverse-biased diode (rectifier), variable resistor (varistor), or RC network, as shown in Fig. 13-10.

The *R* and *C* in an RC circuit are calculated with the following equations:

$$C = \frac{I^2}{10}\ \mu\text{F} \qquad (13\text{-}14)$$

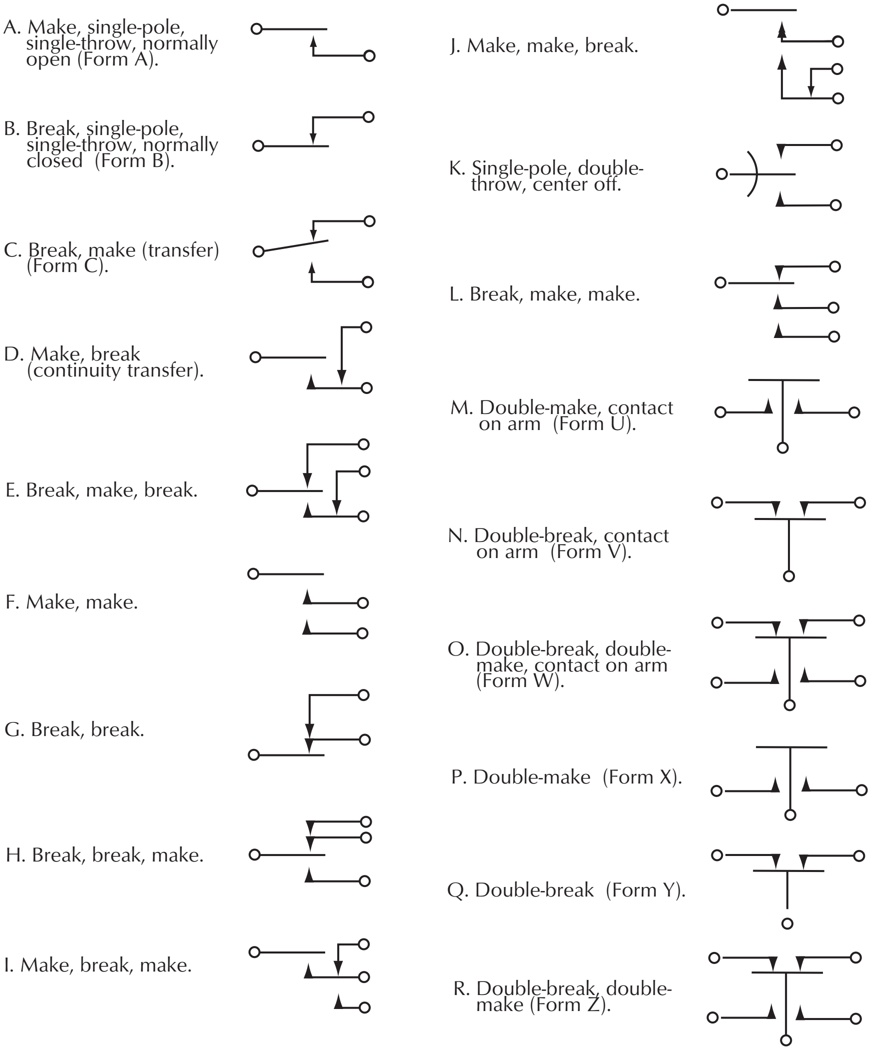

A. Make, single-pole, single-throw, normally open (Form A).
B. Break, single-pole, single-throw, normally closed (Form B).
C. Break, make (transfer) (Form C).
D. Make, break (continuity transfer).
E. Break, make, break.
F. Make, make.
G. Break, break.
H. Break, break, make.
I. Make, break, make.
J. Make, make, break.
K. Single-pole, double-throw, center off.
L. Break, make, make.
M. Double-make, contact on arm (Form U).
N. Double-break, contact on arm (Form V).
O. Double-break, double-make, contact on arm (Form W).
P. Double-make (Form X).
Q. Double-break (Form Y).
R. Double-break, double-make (Form Z).

Figure 13-9. Various contact arrangements of relays. (From *American National Standard Definitions and Terminology for Relays for Electronics Equipment C83.16-1971.*)

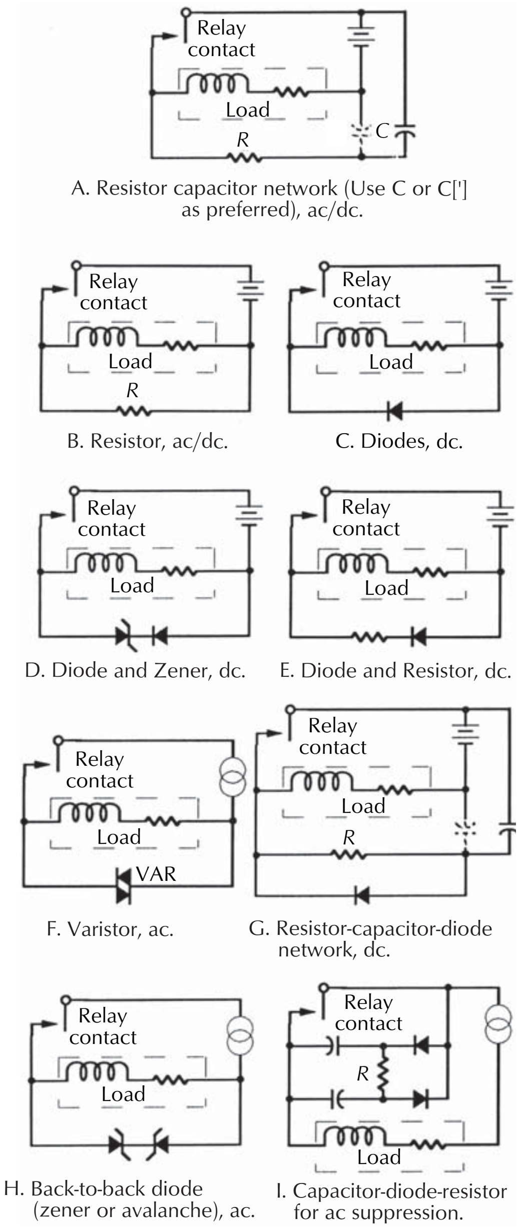
$$R = \frac{0.01\ V}{I^{\left(1+\frac{50}{V}\right)}} \tag{13-15}$$

When using a rectifier, the rectifier is an open circuit to the power source because it is reverse biased; however, when the circuit breaks, the diode conducts. This technique depends on a reverse path for the diode to conduct; otherwise, it will flow through some other part of the circuit. It is important that the rectifier have a voltage rating equal to the transient voltage.

Contact bounce occurs in all mechanical-type relays except the mercury-wetted types that, because of the thin film of mercury on the contacts, do not break during make. Bounce creates noise in the circuit, particularly when switching audio where it acts as a dropout.

13.2.3 Relay Loads[8]

Never assume that a relay contact can switch its rated current no matter what type of load it sees. High in-rush currents or high induced back electromotive force (emf) like those of Fig. 13-11 can quickly erode or weld elec-



A. Resistor capacitor network (Use C or C['] as preferred), ac/dc.

B. Resistor, ac/dc.

C. Diodes, dc.

D. Diode and Zener, dc.

E. Diode and Resistor, dc.

F. Varistor, ac.

G. Resistor-capacitor-diode network, dc.

H. Back-to-back diode (zener or avalanche), ac.

I. Capacitor-diode-resistor for ac suppression.

Figure 13-10. Methods of suppressing transients across contacts. Courtesy Magnecraft Electric Co.

tromechanical relay contacts and destroy solid-state relays.

13.2.3.1 The Effects of Various Loads

Incandescent Lamps. The cold resistance of a tungsten-filament lamp is extremely low, resulting in in-rush currents as much as 15 times the steady-state current. This is why lamp burnout almost always occurs during turn on.

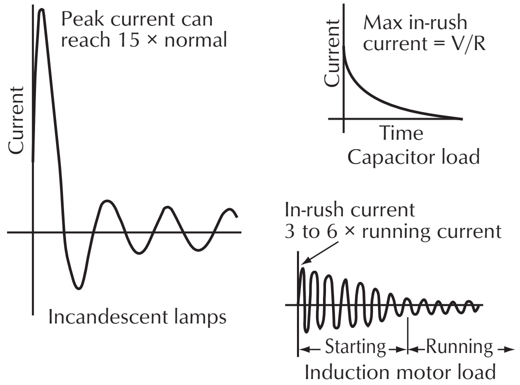


Figure 13-11. High in-rush current on turn-on can damage relays.

Capacitive Loads. The initial charging current to a capacitive circuit can be extremely high, since the capacitor acts as a short circuit, and current is limited only by the circuit resistance. Capacitive loads may be long transmission lines, filters for electromagnetic interference (emi) elimination, and power supplies.

Motor Loads. High in-rush current is drawn by most motors, because at standstill their input impedance is very low. This is particularly bad when aggravated by contact bounce causing several high-current makes and breaks before final closure. When the motor rotates, it develops an internal back emf that reduces the current. Depending on the mechanical load, the starting time may be very long and produce a relay-damaging in-rush current.

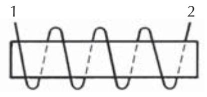
Inductive Loads. In-rush current is limited by inductance; however, when turned off, energy stored in magnetic fields must be dissipated.

Dc Loads. These are harder to turn off than ac loads because the voltage never passes through zero. When electromagnetic radiation (emr) contacts open, an arc is struck that may be sustained by the applied voltage, burning contacts.

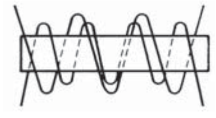
13.2.4 Electromechanical Relays

Regardless of whether the relay operates on ac or dc, it will consist of an actuating coil, a core, an armature, and a group of contact springs that are connected to the circuit or circuits to be controlled. Associated with the armature are mechanical adjustments and springs. The mechanical arrangement of the contacts may be such that when the relay is at rest, certain circuits are either open or closed. If the contacts are open when the relay is at rest (not energized) they are called *normally open* contacts.

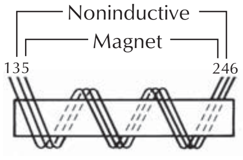
Relays are wound in many different manners, Fig. 13-12. Among them are the single wound, double wound, trifilar wound, bifilar wound, and two coil, which are nonelectromagnetic.



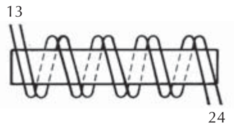
A. Single.



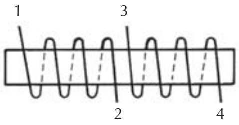
B. Double.



C. Trifilar.



D. Bifilar.



E. Two coil.

Figure 13-12. Types of relay coil windings.

13.2.4.1 Dc Relays

Direct current (dc) relays are designed to operate at various voltages and currents by varying the dc resistance of the actuating coils, and may vary from a few to several thousand ohms. Dc relays may operate as marginal, quick-operate, slow-operate, or polarized.

A *marginal relay* operates when the current through its winding reaches a specified value, and it releases when the current falls to a given value.

In the *quick-operate* type, the armature is attracted immediately to the pole piece of the electromagnet when the control circuit is closed.

Slow-operate relays have a time-delay characteristic; that is, the armature is not immediately attracted to the pole piece of the electromagnet when the control circuit is closed. To accomplish this a copper collar is placed around the armature end of the pole piece. They differ from the slow-release variety in that the latter type has the copper collar around the end of the pole piece opposite from the armature.

A *polarized relay* is designed to react to a given direction of current and magnitude. Polarized relays use a permanent magnet core. Current in a given direction increases the magnetic field, and in the opposite direction it decreases the field. Thus, the relay will operate only for a given direction of current through the coil.

A *latching relay* is stable in both positions. One type of latching relay contains two separate actuating coils. Actuating one coil latches the relay in one position where it remains until it is unlatched by energizing the other coil.

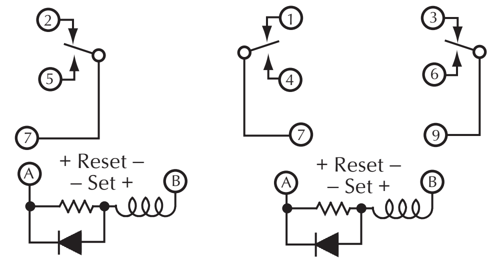
A second and more modern type is a bistable magnetic latching relay. This type is available in single- or dual-coil latching configurations. Both are bistable and will remain in either state indefinitely. The coils are designed for intermittent duty: 10 s maximum on-time. The relay sets or resets on a pulse of 100 ms or greater. Fig. 13-13 shows the various contact and coil forms.

13.2.4.2 Ac Relays

Alternating-current (ac) relays are similar in construction to the dc relays. Since ac has a zero value every half cycle, the magnetic field of an ac-operated relay will have corresponding zero values in the magnetic field every half cycle.

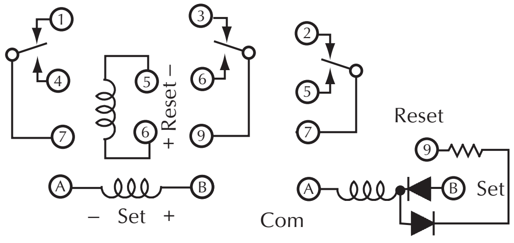
At and near the instants of zero current, the armature will leave the core, unless some provision is made to hold it in position. One method consists of using an armature of such mass that its inertia will hold it in position. Another method makes use of two windings on separate cores. These windings are connected so that their respective currents are out of phase with each other. Both coils effect a pull on the armature when current flows in both windings.

A third type employs a split pole piece of which one part is surrounded by a copper ring acting as a shorted turn. Alternating current in the actuating coil winding induces a current in the copper coil. This current is out of phase with the current in the actuating coil and does not reach the zero value at the same instant as the



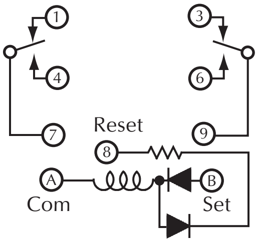
A. Dc single coil, 1 Form C contact

B. Dc single coil, 2 Form C contacts.



C. Dc dual coil, 2 Form C contacts.

D. Ac coil, 1 Form C contact.



E. Ac coil, 2 Form C contacts.

Figure 13-13. Various types and pin connections for latching relays. Courtesy Magnecraft Electric Co.

current in the actuating coil. As a result, there is always enough pull on the armature to hold it in the operating position.

An *ac differential relay* employs two windings exactly alike, except they are wound in opposite directions. Such relays operate only when one winding is energized. When both windings are energized in opposite directions, they produce an aiding magnetic field, since the windings are in opposite directions. When the current through the actuating coils is going in the same direction, the coils produce opposite magnetic fields. If the current through the two coils is equal, the magnetic fields neutralize each other and the relay is nonoperative.

A differential polar relay employs a split magnetic circuit consisting of two windings on a permanent magnet core. A differential polar relay is a combination of a differential and a polarized relay.

13.2.5 Reed Relays[6,7,8,11]

Reed relays were developed by the Bell Telephone Laboratories in 1960 for use in the Bell System central offices. The glass envelope is surrounded by an electromagnetic coil connected to a control circuit. Although originally developed for the telephone company, such devices have found many uses in the electronics industry.

The term *reed relay* covers dry reed relays and mercury-wetted contact relays, all of which use hermetically sealed reed switches. In both types, the reeds (thin, flat blades) serve multiple functions, as conductor, contacts, springs, and magnetic armatures. Reed relays are usually soldered directly onto a circuit board or plugged into a socket that is mounted onto a circuit board.

13.2.5.1 Contact Resistance and Dynamics

Reed relays have much better switching speed than electromechanical relays. The fastest Coto Technology switching reed relay is the 9800 series, with a typical actuate time of 100 µs. Release time is approximately 50 µs. Actuate time is defined as the period from coil energization until the contact is closed and has stopped bouncing. After the contacts have stopped bouncing, they continue to vibrate while in contact with one another for a period of about 1 ms. This vibration creates a wiping action and variable contact pressure.

Static contact resistance (SCR) is the resistance across the contact terminals of the relay after it has been closed for a sufficient period of time to allow for complete settling. For most reed relays, a few milliseconds is more than adequate, but the relay industry uses 50 ms to define the measurement.

Another contact resistance measurement that has provided great insight into the overall quality of the relay is contact resistance stability (CRS). CRS measures the repeatability of successive static contact resistance measurements.

13.2.5.2 Magnetic Interaction

Reed relays are subject to external magnetic effects including the earth's magnetic field (equivalent to approximately 0.5 AT and generally negligible), electric motors, transformers external magnets, etc., which may change performance characteristics. Such magnetic sources include one common source of an external magnetic field acting on a relay or another relay operating in close proximity. The potential for magnetic coupling must be taken into account when installing densely packed single- or multichannel relays.

An example of magnetic interaction is shown in Fig. 13-14 where two relays, K1 and K2, with identical coil polarities are mounted adjacent to each other. When K2 is "off", relay K1 operates at its designed voltage. When K2 is activated, the magnetic fields oppose so the effective magnetic flux within K1 is reduced, requiring an increase in coil voltage to operate the reed switch. For closely packed relays without magnetic shields, a 10–20% increase in operate voltage is typical, which can drive the relays above their specified limits. The opposite effect occurs if K1 and K2 are polarized in opposite directions making the operating voltage for K1 less.

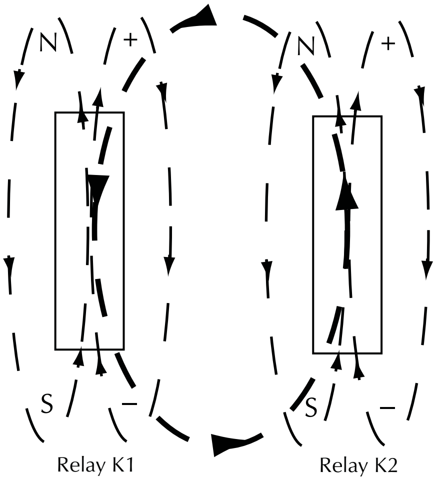


Figure 13-14. Adverse magnetic interaction. Courtesy Coto Technology.

There are several ways to reduce magnetic interaction between relays:

- Specify relays that incorporate an internal or external magnetic shield.
- Apply an external magnetic shield to the area where the relays are mounted. A sheet of mu-metal or other high-magnetic-permeability ferrous alloy 2–5 mils thick is effective.
- Provide increased on-center spacing between relays. Each doubling of this distance reduces the interaction effect by a factor of approximately four.
- Avoid simultaneous operation of adjacent relays.
- Provide alternating coil polarities for relays used in a matrix.

13.2.5.3 Environmental Temperature Effects

The resistance of the copper wire used in reed relay coils increases by 0.4% /1°C rise in temperature. Reed relays are current-sensitive devices so their operate and release levels are based on the current input to the coil. If a voltage source is used to drive the relays, an increase in coil resistance causes less current to flow through the coil, so the voltage must be increased to compensate and maintain current flow. Industry standards define that relays are typically specified at 25°C ambient. If the relay is used in higher ambient conditions or near external sources of heat, this must be carefully considered.

For example, a standard relay nominally rated at 5 Vdc has a 3.8 Vdc maximum operate value at 25°C as allowed by the specifications. If the relay is used in a 75°C environment, the 50°C temperature rise increases the operate voltage by 50 × 0.4%, or 20%. The relay now will operate at 3.8 Vdc + (3.8 Vdc × 20%), or 4.56 Vdc. If there is more than a 0.5 Vdc drop in supply voltage due to a device driver or sagging power supply, the relay may not operate. Under these conditions there will be increases in operate and release timing to approximately the same 20%.

13.2.5.4 Dry Reed Relays

Because of the tremendous increases in low-level logic switching, computer applications, and other business machine and communication applications, dry reed relays have become an important factor in the relay field. They have the great advantage of being hermetically sealed, making them impervious to atmospheric contamination. They are very fast in operation and when operated within their rated contact loads, they have a very long life. They can be manufactured automatically and therefore are relatively inexpensive. A typical dry reed switch capsule is shown in Fig. 13-15.

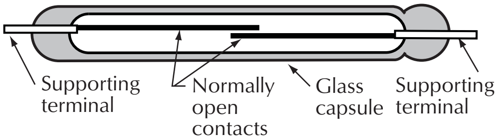


Figure 13-15. Construction of a switch capsule of a typical dry reed relay—Form A. Courtesy Magnecraft Electric Co.

In this basic design, two opposing reeds are sealed into a narrow glass capsule and overlap at their free ends. At the contact area, they are plated with rhodium over gold to produce a low contact resistance when they meet. The capsule, surrounded by an electromagnetic coil, is made of glass and filled with a dry inert gas. When the coil is energized in the basic Form A contact combination, the normally open contacts are brought together; when the field is removed the reeds separate by their own spring tension.

Some may contain permanent magnets for magnetic biasing to achieve normally closed contacts (Form B). Single-pole, double-throw contact combinations (Form C) are also available. Current rating, which is dependent on the size of the reed and the type and amount of plating, may range from low level to 1 A. Effective contact protection is essential in most applications unless switching is done dry.

Relay packages using up to four Form C and six Form A dry reed switches are common, providing multiple switching arrangements. The reed relay may be built for a large variety of operational modes such as pulse relay, latch relay, crosspoint relay, and logic relay. These relays may also be supplied with electrostatic or magnetic shields. The relay in Fig. 13-16 has two Form C contacts.

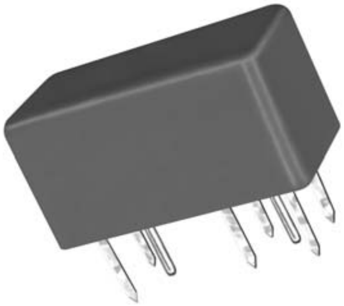


Figure 13-16. Coto Technology 2342 multipole relay, Courtesy Coto Technology.

Reed switches have the following characteristics:

- A high degree of reliability stemming from their controlled contact environment.
- Consistency of performance resulting from a minimum number of parts.
- Long operational life.
- Ease of packaging as a relay.
- High-speed operation.
- Small size.
- Low cost.

Number of Switches. There appears to be no limit to the number of switches that can be actuated by a common coil. However, as the number increases, coil efficiency decreases and power input increases. This can lead to a practical limitation. On the other hand, the increase in power required to operate one more switch capsule is usually less than the total required if the assembly were split in two. The single contact relay is the most frequently used but relays with four or more switches in a single coil are quite common.

Sensitivity. The power input required to operate dry reed relays is determined by the sensitivity of the particular reed switch used, by the number of switches operated by the coil, by the permanent magnet biasing (if used), and by the efficiency of the coil and the effectiveness of its coupling to the reeds. The minimum input required to effect closure ranges from milliwatts for a single capsule sensitive unit to several watts for a multipole relay.

Operate Time. Coil time constant, overdrive, and the characteristics of the reed switch determine operate time. With maximum overdrive, reed relays will operate in approximately 200 μs or less. Drive at rated voltage usually results in a 1 ms operate time.

Release Time. With the relay coil unsuppressed, dry reed switch contacts release in a fraction of a millisecond. Form A contacts open in as little as 50 μs. Magnetically biased Form B contacts and normally closed contacts of Form C switches reclose from 100 μs to 1 ms, respectively.

If the relay coil is suppressed, release times are increased. Diode suppression can delay release for several milliseconds, depending on coil characteristics, drive level, and reed release characteristics.

Bounce. As with the other hard contact switches, dry reed contacts bounce on closure. The duration of bounce is typically quite short and is in part dependent on drive level. In some of the faster devices, the sum of operate time and bounce is relatively constant so as drive is increased, the operate time decreases and bounce increases.

While normally closed contacts of a Form C switch bounce more than normally open contacts, magnetically biased Form B contacts exhibit essentially the same bounce as Form A.

Contact Resistance. Because the reeds in a dry reed switch are made of a magnetic material that has a high volume resistivity, terminal-to-terminal resistance is somewhat higher than in some other types of relays. Typical specification limit for initial maximum resistance of a Form A reed relay is 0.200 Ω.

13.2.5.5 Mercury-Wetted Contact Relays

Mercury-wetted contact relays are a form of reed relays consisting of a glass-encapsulated reed with its base immersed in a pool of mercury and the other end capable of moving between one or two stationary contacts. The mercury flows up to the reed by capillary action and wets the contact surface of the moving end of the reed as well as the contact surfaces of the stationary contacts. Thus a mercury-to-mercury contact is maintained in a closed position. The mercury-wetted relay is usually actuated by a coil around the capsule.

Aside from being extremely fast in operation and having relatively good load-carrying capacity, mercury-wetted contact relays have extremely long life since the mercury films are reestablished at each contact closure and contact erosion is eliminated. Since the films are "stretchable," there is no contact bounce. Contact interface resistance is extremely low.

Three disadvantages of this type of reed relays are:

1. The freezing point of mercury is (–38.8°C or –37.8°F).
2. They have poor resistance to shock and vibration.
3. Some type need to mount in a near vertical position.

These relays are available in a compact form for printed-circuit board mounting. Multipole versions can be provided by putting additional capsules inside the coil. They are used for a great variety of switching applications such as are found in computers, business machines, machine tool control systems, and laboratory instruments.

Mercury-wetted switches can also come as a nonposition sensitive, mercury-wetted, reed relay that combines the desirable features of both dry reed and mercury-wetted capsules. This allows the user to place the reed relay in any position and is capable of withstanding shock and vibration limits usually associated with dry reed capsules. On the other hand, they retain the principal advantages of other mercury-wetted switches—no contact bounce and low stable contact resistance.

Operation of the nonposition-sensitive switch is made possible by the elimination of the pool of mercury at the bottom of the capsule. Its design captures and retains the mercury on contact and blade surfaces only. Due to the limited amount of mercury film, this switch should be restricted for use at low-level loads.

Mercury-wetted reed relays are a distinct segment of the reed relay family. They are different from the dry reed relays in the fact that contact between switch elements is made via a thin film of mercury. Thus, the most important special characteristics of mercury-wetted relays are:

- Contact resistance is essentially constant from operation to operation throughout life.
- Contacts do not exhibit bounce. The amount of mercury at the contacts is great enough to both cushion the impact of the underlying members and to electrically bridge any mechanical bounce that remains.
- Life is measured in billions of operations, due to constant contact surface renewal.
- Contacts are versatile. The same contacts, properly applied, can handle relatively high-power and low-level signals.
- Electrical parameters are constant. With contact wear eliminated, operating characteristics remain the same through billions of operations.

To preserve these characteristics, the rate of change of voltage across the contacts as they open must be limited to preclude damage to the contact surface under the mercury. For this reason, suppression should be specified for all but low-level applications.

Mounting Position. To ensure that distribution of mercury to the relay contacts is proper, position sensitive types should be mounted with switches oriented vertically. It is generally agreed that deviation from vertical by as much as 30° will have some effect on performance. The nonposition-sensitive mercury-wetted relay, which is the most common type today, is not affected by these limitations.

Bounce. Mercury-wetted relays do not bounce if operated within appropriate limits. However, if drive rates are increased, resonant effects in the switch may cause rebound to exceed the level that can be bridged by the mercury, and electrical bounce will result. Altered distribution of mercury to the contacts, caused by the high rate of operation, may also contribute to this effect.

Contact Resistance. Mercury-wetted relays have a terminal-to-terminal contact resistance that is somewhat lower than dry reed relays. Typical specification limit for maximum contact resistance is 0.150 Ω.

13.2.5.6 RF Relays

RF relays are used in high-frequency applications, usually in a 50 Ω circuit. The RF coaxial shielded relay in Fig. 13-17 can switch up to 200 Vdc at 0.5 A.

Figure 13-17. Coto Technology 9290 RF reed relay. Courtesy Coto Technology.

Insertion and Other Losses. In the past, the typical parameters used to quantify RF performance of reed relays were Insertion loss, isolation, and return loss (sometimes called reflection loss). These are frequency-related vector quantities describing the relative amount of RF power entering the relay and either being transmitted to the output or being reflected back to the source. For example, with the relay's reed switch closed and 50% power being transmitted through the relay, the insertion loss would be 0.5 or −3 dB. The frequency at which a −3 dB rolloff occurs is a convenient scalar (single-valued) quantity for describing insertion loss performance.

Isolation. The RF isolation of the reed relay can be determined by injecting an RF signal of known power amplitude with the reed switch open (coil unactivated). Sweeping the RF frequency and plotting the amount of RF energy exiting the relay allows the isolation curve to be plotted on a dB scale. At lower frequencies, the isolation may be −40 dB or greater, indicating that less than 0.01% of the incident power is leaking through the relay. The isolation decreases at higher frequencies, because of capacitive leakage across reed switch contacts.

Return Loss. Return loss represents the amount of RF power being reflected back to the source with the reed switch closed and the output terminated with a standard impedance, normally 50 Ω. If the relay was closely matched to 50 Ω at all frequencies, the reflected energy would be a very small fraction of the incident energy from low to high frequencies. In practice, return loss increases (more power is reflected) as frequency increases. High return loss (low reflective energy) is desirable for high-speed pulse transmission, since there is less risk of echoing signal collisions that can cause binary data corruption and increased bit error rates. Return loss is calculated from the reflection coefficient (ρ), which is the ratio of the magnitude of signal power being reflected from a closed relay to the power input at a specified frequency

$$Return\ loss = -20\log\rho \tag{13-16}$$

To determine the RF performance of a reed relay involves injecting a swept frequency RF signal of known power into the relay and measuring the amount of RF energy transmitted through or reflected back from it. These measurements can be conveniently made using a Vector Network Analyzer (VNA). These test instruments comprise a unified RF sweep frequency generator and quantitative receiver/detector. In the case of a Form A relay, the device is treated as a network with one input and one output port, and the amount of RF energy entering and being reflected from each port is recorded as a function of frequency. Thus a complete characterization of a Form A relay comprises four data vectors, designated as follows:

S_{11}	power reflected from input port.
S_{12}	power transmitted to input port from output port.
S_{21}	power transmitted to output port from input port.
S_{22}	power reflected from output port.

Voltage Standing Wave Ratio (VSWR). VSWR is a measurement of how much incident signal power is reflected back to the source when an RF signal is injected into a closed relay terminated with a 50 Ω impedance. It represents the ratio of the maximum amplitude of the reflected signal envelope amplitude divided by the minimum at a specified frequency. A VSWR of 1 indicates a perfect match between the source, relay, and output load impedance and is not achievable. VSWR at any particular frequency can be converted from y-axis return loss using Table 13-2.

Table 13-4. Return Loss Versus VSWR

Return Loss VSWR (dB)	VSWR
–50	1.01
–40	1.02
–30	1.07
–20	1.22
–10	1.93
–3	5.85

Rise Time. The rise time of a reed relay is the time required for its output signal to rise from 10% to 90% of its final value, when the input is changed abruptly by a step function signal. The relay can be approximated by a simple first-order low-pass filter. The rise time is approximately

$$T_r = RC \times \ln\frac{90\%}{10\%} \quad (13\text{-}17)$$
$$= 2.3RC$$

Substituting into the equation for the 50% roll-off frequency $f_{-3\ \mathrm{dB}} = 1/2\pi RC$ yields the relationship

$$T_r = \frac{0.35}{f_{-3\ \mathrm{dB}}}. \quad (13\text{-}18)$$

Therefore the relay's rise time can be simply estimated from the S_{21} insertion loss curve by dividing the –3 dB rolloff frequency into 0.35. For example, the Coto Technology B40 ball grid relay has $f_{-3\ \mathrm{dB}}$ = 11.5 GHz, from which the rise time can be estimated as 30 ps.

Effect of Lead Form on High Frequency Performance. Surface mount (SMD) relays give better RF performance than those with through hole leads. SMD leadforms comprise gullwing, J-bend, and axial forms. Each has its advantages and disadvantages, but the RF performance point of view, axial relays generally have the best RF performance in terms of signal losses, followed by J-bend and gullwing. The straight-through signal path of axial relays minimizes capacitive and inductive reactance in the leads and minimizes impedance discontinuities in the relay, resulting in the highest bandwidth. However, the axial leadform requires a cavity in the printed circuit board to receive the body of the relay. An advantage is the effective reduced height of the axial relay, where space is at a premium.

J-bend relays provide the next-best RF performance and have the advantages of requiring slightly less area on the PCB. The gullwing form is the most common type of SMD relay. It has the longest lead length between the connection to the PCB pad and the relay body which results in slightly lower RF performance than the other lead types. Initial pick-and-place soldering is simple, as is rework, resulting in a broad preference for this lead type unless RF performance is critical.

Coto Technology's new leadless relays have greatly enhanced RF performance. They do not have traditional exposed metal leads; instead, the connection to the user's circuit board is made with ball-grid-array (BGA) attachment, so that the devices are essentially leadless. In the BGA relays, the signal path between the BGA signal input and output is designed as an RF transmission line, with an RF impedance close to 50 Ω throughout the relay. This is achieved using a matched combination of coplanar waveguide and coaxial structures with very little impedance discontinuity through the relays. The Coto B10 and B40 reed relays, Fig. 13-18 achieve bandwidths greater than 10 GHz and rise times of 35 ps or less.

Figure 13-18. Coto Technology B40 Ball Grid surface mount 4-channel reed relay. Courtesy Coto Technology.

Skin Effect in Reed Relays. At high frequencies, RF signals tend to travel near the surface of conductors rather than through the bulk of the material. The skin effect is exaggerated in metals with high magnetic permeability, such as the nickel-iron alloy used for reed switch blades. In a reed switch, the same metal has to carry the switched current and also respond to a magnetic closure field. Skin effect does not appreciably affect the operation of reed relays at RF frequencies because the increase in ac resistance due to skin effect is proportional to the square root of frequency, whereas the losses due to increasing reactance are directly proportional to *L* and inversely proportional to *C*. Also the external lead surfaces are coated with tin or solder alloys for enhanced solder-ability which helps to reduce skin effect losses.

Selecting Reed Relays for High Frequency Service. High-speed switching circuits can be accomplished with reed relays, electromechanical relays (EMRs) specifically designed for high-frequency service, solid-state relays (SSRs), PIN diodes, and microelectromechanical systems (MEMS) relays. In many cases, reed relays are an excellent choice, particularly with respect to their unrivalled RC product. RC is a figure of merit expressed in pF·Ω, where *R* is the closed contact resistance and *C* is the open contact capacitance. The lower this figure is, the better the high-frequency performance. The RC product of a Coto Technology B40 relay for example, is approximately 0.02 pF•Ω. SSRs have pF•Ω products equal to about 6, almost 300 times higher, plus, the breakdown voltage at these pF•Ω levels is much lower than that of a reed switch. The turn-off time for SSRs is also longer than the 50 µs needed by a reed relay to reach its typical 10^{12} Ω off resistance. Some feel that the reliability of reed relays compared to solid-state devices is largely unjustified, due to continuous technological improvements. Many reed relays have demonstrated MCBF values of several hundred million to several billion closure cycles at typical signal switching levels.

PIN diodes are occasionally used for HF switching. However, PIN diodes require relatively complex drive circuitry compared to the simple logic circuitry that drives reed relays. PIN diodes typically have a lower frequency cut-on of about 1 MHz, while a reed relay can switch from dc to its useful cut-off frequency. The high junction capacitance of PIN diodes results in lower RF isolation than a reed relay when the PIN diode is biased open. When biased closed, the higher on-resistance of the PIN diode can lead to Q-factor damping in the circuit to which it is connected. PIN diodes can exhibit significant nonlinearity, leading to gain compression, harmonic distortion, and intermodulation distortion, while reed relays are linear switching devices.

Electromechanical relays (EMRs) have been developed with bandwidths to about 6 GHz, and isolation of about −20 dB at that frequency. This isolation is better than that of a reed relay, since the contacts can be designed with bigger spacing, resulting in lower capacitive leakage. This advantage must be weighed against the increased size and cost of EMRs and lower reliability. The EMR has a complex structure with more moving parts than the simple blade flexure involved in closing a reed switch, resulting in a lower mechanical life. If higher isolation is required with a reed relay solution, two relays can be cascaded together with a combined reliability that is still higher than that of a typical EMR.

MEMS switches (relays) are being developed based on two technologies, electrostatic closure and pulsed magnetic toggling between open and closed states. They offer potential advantages in terms of small and low loss high-frequency switching. However, adequate contact reliability has not been demonstrated at the switching loads required by automated test equipment (ATE) applications. At present, though, MEMS relay technology is too immature for use in most applications addressed by reed relays.

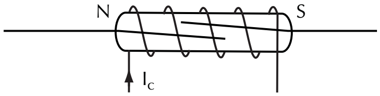
13.2.5.7 Dry Reed Switches

A dry reed switch is an assembly containing ferromagnetic contact blades that are hermetically sealed in a glass envelope and are operated by an externally generated magnetic field. The field can be a coil or a permanent magnet. The switches in Figs. 13-19A and 13-19B can switch up to 175 Vdc at 350 mA or 140 Vac at 250 ma. The switch in Fig. 13-19C can switch 200 Vdc at 1 A or 140 Vac at 1 A.

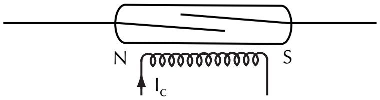
A. RI-80 SMD SPST 5 W switch. B. RI-80 SPST 5W switch.

C. RI-25 SPST 25 W switch.

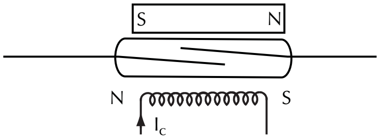
Figure 13-19. Coto Technology dry reed switches. Courtesy Coto Technology.



A. A dry-reed switch mounted within a coil.

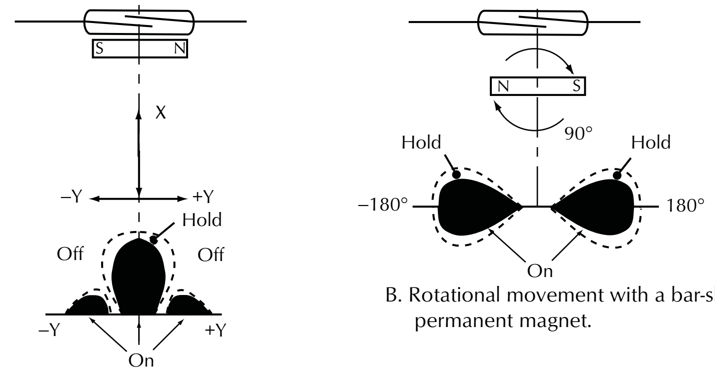


B. A dry-reed switch mounted outside a coil.



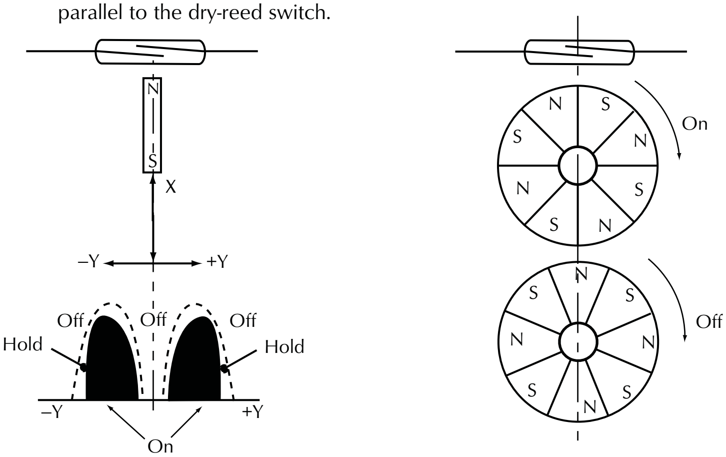
C. A dry-reed switch biased by a permanent magnet and operated by a coil.

Figure 13-20. Energizing a dry-reed switch with a coil. Courtesy Coto Technology.



A. Movement with the magnetic field parallel to the dry-reed switch.

B. Rotational movement with a bar-shaped permanent magnet.



C. Movement with the magnetic field perpendicular to the dry-reed switch.

D. Rotational movement with two or more ring magnets.

Figure 13-21. Energizing a dry-reed switch with a permanent magnet. Courtesy Coto Technology.

Fig. 13-20 shows three methods of operating a reed switch using a coil. Fig. 13-21 shows four ways to operate a reed switch using permanent magnets.

13.2.6 Solid-State Relays[9]

Solid-state relays (SSRs) utilize the on–off switching properties of transistors and SCRs for opening and closing dc circuits. They also use triacs for switching ac circuits.

13.2.6.1 Advantages

SSRs have several advantages over their electromechanical counterparts: no moving parts, arcing, burning, or wearing of contacts; and the capacity for high-speed, bounceless, noiseless operation. Many SSRs are available that feature optical coupling; thus, the signal circuit includes a lamp or light-emitting diode that shines on a phototransistor serving as the actuating device. In other types of SSRs, a small reed relay or transformer may serve as the actuating device. A third type is direct coupled and therefore not actually an SSR because there is no isolation between input and output. These are better called an amplifier. All three types are shown in Fig. 13-22.

Ac relays turn on and off at zero crossing; therefore, they have reduced dv/dt. However, this does slow down the action to the operating frequency.

13.2.6.2 Disadvantages and Protection[4]

Solid-state relays also have some inherent problems as they are easily destroyed by short circuits, high surge current, high dv/dt, and high peak voltage across the power circuit.

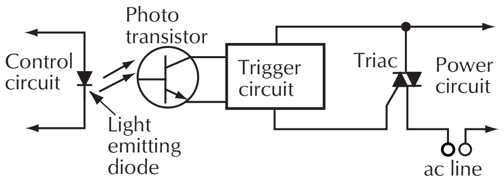
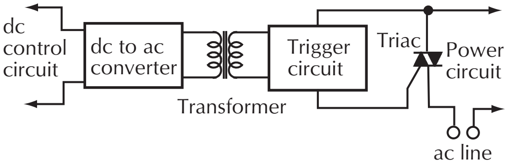
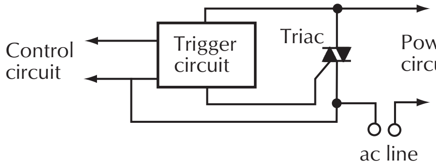


Photo
transistor
Control
circuit
Trigger
circuit
Triac
Power
circuit
Light
emitting
diode
ac line

A. Optically coupled.


dc
control
circuit
dc to ac
converter
Trigger
circuit
Triac
Power
circuit
Transformer
ac line

B. Transformer coupled.


Control
circuit
Trigger
circuit
Triac
Power
circuit
ac line

C. Direct coupled.

Figure 13-22. Various types of solid-state relays.

Short-circuit and high-surge current protection is performed with fast blow fuses or series resistors. A standard fuse normally will not blow before the SCR or triac is destroyed since the fuses are designed to withstand surge currents. Fast blow fuses will act on high in-rush currents and usually protect solid-state devices.

Using a current-limiting resistor will protect the SSR; however, it creates a voltage drop that is current dependent and, at high current, dissipates high power.

A common technique for protecting solid-state switching elements against high *dv*/*dt* transients is by shunting the switching element with an RC network (snubber), as shown in Fig. 13-23. The following equations provide effective results:

$$R_1 = \frac{L}{V} \times \frac{dv}{dt} \tag{13-19}$$

$$R_2 = \frac{\sqrt{1-(PF)^2}}{2\pi f} \times \frac{dv}{dt} \tag{13-20}$$

$$C = \frac{4L}{R_2^2} \tag{13-21}$$

$$C = \frac{4}{R_2{}^2} \times \frac{V}{I} \times \frac{\sqrt{1-PF^2}}{2\pi F} \tag{13-22}$$

where,

L is the inductance in henrys,

V is the line voltage,

dv/dt is the maximum permissible rate of change of voltage in volts per microsecond,

I is the load current,

PF is the load power factor,

C is the capacitance in microfarads,

R_1, R_2 are the resistance in ohms,

f is the line frequency.

RC networks are often internal to SSRs.

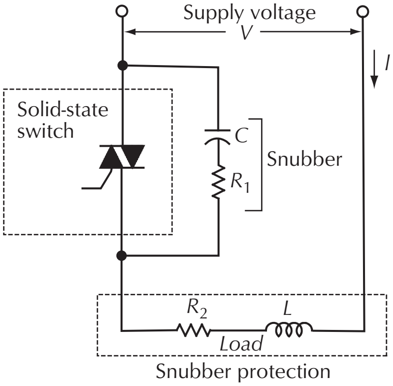

Supply voltage
V
I
Solid-state
switch
C
R1
Snubber
R2
L
Load
Snubber protection

Figure 13-23. Snubber circuit for solid-state relay protection.

13.2.6.3 High-Peak-Transient-Voltage Protection

Where high-peak-voltage transients occur, effective protection can be obtained by using metal-oxide varistors (MOVs). The MOV is a bidirectional voltage-sensitive device that becomes low impedance when its design voltage threshold is exceeded.

Fig. 13-24 shows how the proper MOV can be chosen. The peak nonrepetitive voltage (V_{DSM}) of the selected relay is transposed to the MOV plot of peak voltage versus peak amperes. The corresponding current for that peak voltage is read off the chart. Using this value of current (I) in

$$V_{DSM} = V_p - IR \tag{13-23}$$

where,

I is the current,

V_p is the peak instantaneous voltage transient,

R is the load plus source resistance.

It is important that the V_{DSM} peak nonrepetitive voltage of the SSR is not exceeded.

The energy rating of the MOV must not be exceeded by the value of

$$E = V_{DSM} \times I \times t \qquad (13\text{-}24)$$

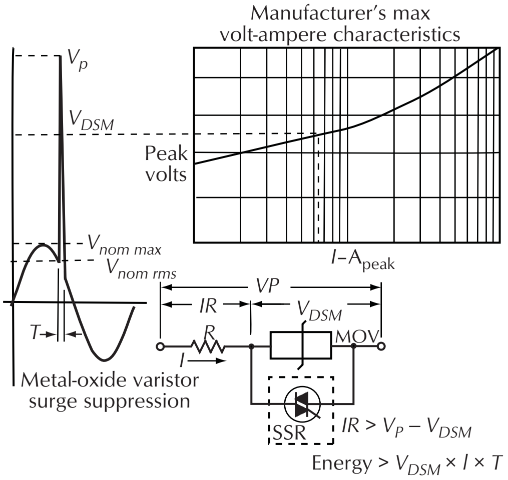

Manufacturer's max volt-ampere characteristics
V_p
V_{DSM}
Peak volts
$V_{nom\ max}$
$V_{nom\ rms}$
I–A_{peak}
VP
IR
V_{DSM}
R
MOV
I
T
Metal-oxide varistor surge suppression
SSR
$IR > V_P - V_{DSM}$
Energy > $V_{DSM} \times I \times T$

Figure 13-24. Metal-oxide varistor peak transient protector.

13.2.6.4 Low Load Current Protection

If the load current is low, it may be necessary to take special precautions to ensure proper operation. Solid-state relays have a finite off-state leakage current. SSRs also need a minimum operating current to latch the output device.

If the off-state voltage across the load is very high, it could cause problems with circuit dropout and component overheating. In these applications a low-wattage incandescent lamp in parallel with the load offers a simple remedy. The nonlinear characteristics of the lamp allow it to be of lower resistance in the off state while conserving power in the on state. It must be remembered to size the SSR for the combined load.

13.2.6.5 Optically Coupled Solid-State Relays

The optically coupled solid-state relay arrangement (SSR) shown in Fig. 13-22A is capable of providing the highest control/power-circuit isolation—many thousands of volts in compact, convenient form. The triac trigger circuit is energized by a phototransistor, a semiconductor device (encapsulated in transparent plastic) whose collector-emitter current is controlled by the amount of light falling on its base region.

A phototransistor is mounted in a light-tight chamber with a light-emitting diode, the separation between them being enough to give high isolation (thousands of volts) between the control and power circuit.

The light-emitting diode requires only 1.5 V to energize and has very rapid response time. The power circuit consists of a high-speed phototransistor and an SCR for dc power source, as well as a triac for ac application.

The relay not only responds with high speed but is also capable of very fast repetitious operation and provides very brief delays in turnoff. In some applications, the photocoupler housing provides a slotted opening between the continuously lit light-emitting diode and the phototransistor. On–off control is provided by a moving arm, vane, or other mechanical device that rides in the slot and interrupts the light beam in accordance with some external mechanical motion. Typical optically coupled SSRs have the following characteristics:

Turn-on control voltage	3–30 Vdc
Isolation	1500 Vac
dv/dt	100 V/µs
Pickup control voltage	3 Vdc
Dropout control voltage	1 Vdc
One-cycle surge (rms)	7–10 times nominal
1 second overload	2.3 times nominal
Maximum contact voltage drop	1.5–4 V

13.2.6.6 Transformer-Coupled Solid-State Relays

In Fig. 13-22B, the dc control signal is changed to ac in a converter circuit, the output of which is magnetically coupled to the triac trigger circuit by means of a transformer. Since there is no direct electrical connection between the primary and secondary of the transformer, control/power-circuit isolation is provided up to the voltage withstanding limit of the primary/secondary insulation.

13.2.6.7 Direct-Coupled Solid-State Relays

The circuit shown in Fig. 13-22C cannot truly be called a solid-state relay because it does not have isolation between input and output. It is the simplest configuration; no coupling device is interposed between the control and actuating circuits, so no isolation of the control circuit is provided. This circuit would be better called an amplifier.

One other variation of these solid-state circuits is occasionally encountered—the *Darlington* circuit. A typical arrangement is shown in Fig. 13-25. Actually a pair of cascaded power transistors, this circuit is used in many solid-state systems to achieve very high power gain—1000 to 10,000 or more. Now marketed in single-transistor cases, it can be obtained as what appears to be a single transistor with high operating voltage ratings that control high amperage loads with only a few volts at the base connection and draw only a few milliamperes from the control circuit. It can be used for relay purposes in a dc circuit the same way, either by direct control signal coupling or with intermediate isolation devices like those described. It is not usable in ac power circuits.

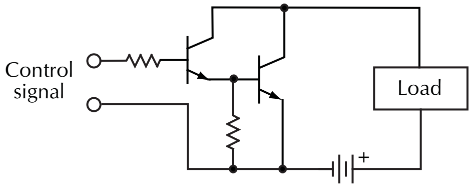

Control
signal
Load
+

Figure 13-25. Darlington direct-coupled solid-state relay.

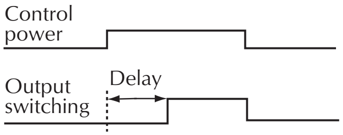
13.2.6.8 Solid-State Time-Delay Relays[10]

Solid-state time-delay relays, Fig. 13-26, can operate in many different modes since they do not rely on heaters or pneumatics. Simple ICs allow the relays to do standard functions plus totaling, intervals, and momentary action as described in the following.

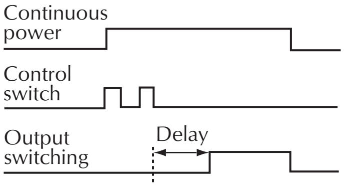
On-Delay. Upon application of control power, the time-delay period begins. At the end of time delay, the output switch operates. When control power is removed, the output switch returns to normal, Fig. 13-26A.

Nontotalizer. Upon the opening of the control switch, the time-delay period begins. However, any control switch closure prior to the end of the time delay will immediately recycle the timer. At the end of the time-delay period, the output switch operates and remains operated until the required continuous power is interrupted, as shown in Fig. 13-26B.

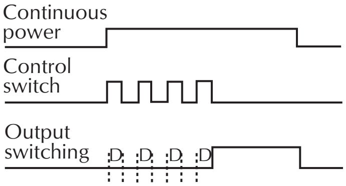
Totalizer/Preset Counter. The output switch will operate when the sum of the individual control switch closure durations equal the preset time-delay period. There may be interruptions between the control switch closures without substantially altering the cumulative timing accuracy. The output switch returns to normal when the continuous power is interrupted, as shown in Fig. 13-26C.


Control
power
Output
switching
Delay

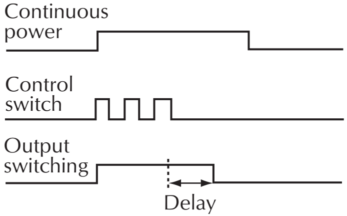
A. On delay.


Continuous
power
Control
switch
Output
switching
Delay

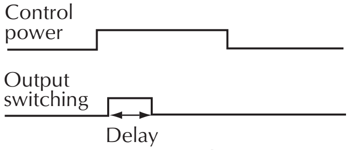
B. Nontotalizer.


Continuous
power
Control
switch
Output
switching
D
D
D
D

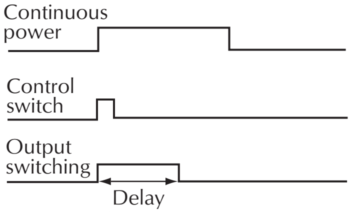
C. Totalizer/preset counter.


Continuous
power
Control
switch
Output
switching
Delay

D. Off delay.


Control
power
Output
switching
Delay

E. Interval.


Continuous
power
Control
switch
Output
switching
Delay

F. Momentary actuation.

Figure 13-26. Types of time-delay relays.

Off-Delay. Upon closure of the control switch, the output switch operates. Upon opening of the control switch, the time-delay period begins. However, any control switch closure prior to the end of the time-delay period will immediately recycle the timer. At the end of the time-delay period, the output switch returns to normal. Continuous power must be furnished to this timer, as shown in Fig. 13-26D.

Interval. Upon application of the control power, the output switch operates. At the end of the time-delay period, the output switch returns to normal. Control power must be interrupted in order to recycle, as shown in Fig. 13-26E.

Momentary Actuation. Upon closure of the control switch, the output switch operates, and the time-delay period begins. The time-delay period is not affected by duration of the control switch closure. At the end of the time-delay period, the output switch returns to normal. Continuous power must be furnished to this timer, as shown in Fig. 13-26F.

Programmable Time-Delay Relay. Programmable time-delay relays are available where the time and functions can be programmed by the user. The Magnecraft W211PROGX-1 relay in Fig. 13-27 is an example of this type. It plugs into an octal socket, has ±0.1% repeatability and four input voltage ranges. It has four programmable functions, On Delay, Off Delay, One Shot, and On Delay and Off Delay. There are 62 programmable timing ranges from 0.1 s to 120 min and the relay has 10 A DPDT contacts. An eight position DIP switch is used to program the timing function and a calibrated knob is used to set the timing.

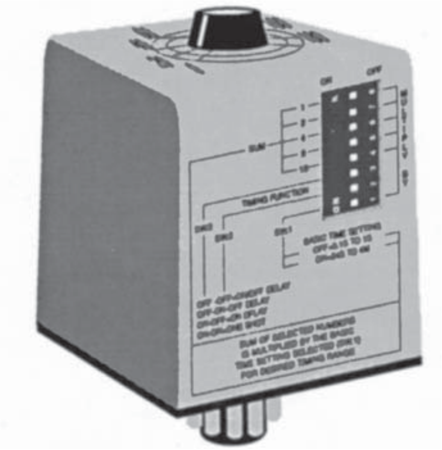


Figure 13-27. A programmable time delay relay. Courtesy Magnecraft Electric Co.

References

1. Logan Hedin, TenHex and Gary Arnold, *Thermal Design of an LED Commercial Florescent Tube,* Cool Polymers; Presented at IMAPs New England 2008 35th Symposium.
2. Solver-Calc©, An Excel Spreadsheet Program for Rapid Thermal Management Calculations, by Villaume Associates, LLC, www.thermalseminars.com.
3. Heat pipes have become increasingly useful in spreading heat loads and moving heat to an area where the convective fin surface area may be more effectively positioned in the available air flow path. “Google” for vendors as there are quite a few suppliers of this component in a thermal solution.
4. TPG is a product designator for highly oriented pyrolytic graphite offered by Momentive Performance Materials in Ohio. The in plane thermal conductivity is ~1750 W/m°K. It can perform in high watt density (> 80 W/in^2) applications without suffering “burn-out” failures that are possible with regular heat pipes.
5. Two major suppliers; Wakefield Thermal Solutions, Pelham, N.H., and Cool Options Inc., Warwick, R.I.
6. *Designer’s Handbook and Catalog of Reed and Mercury Wetted Contact Relays*, Chicago: Magnecraft Electric Co.
7. “Relays, Solenoids, Timers, and Motors,” *Electronic Buyer's Handbook*, 1978. Copyright 1978 by CMP Publications. Reprinted with permission.
8. “Relay Loads” *Electronic Design* 26, December 20, 1978.
9. *Solid-State Relays*, Chicago: Magnecraft Electric Co.
10. *Designer's Handbook and Catalog of Time-Delay Relays,* Chicago: Magnecraft Electric Co.
11. *Coto Technology Reed Relays and Dry Reed Switches*, Coto Technology.

Chapter 14

Transmission Techniques: Wire and Cable

by Steve Lampen and Glen Ballou

14.1 Introduction 397
14.2 Conductors 397
  14.2.1 Resistance and Wire Size 397
  14.2.2 Calculating Wire Resistance 397
  14.2.3 Resistance and Gage Size 398
  14.2.4 Drawing and Annealing. 399
  14.2.5 Plating and Tinning 399
  14.2.6 Conductor Parameters 399
    14.2.6.1 Stranded Cables. 400
  14.2.7 Pulling Tension 401
  14.2.8 Skin Effect. 402
  14.2.9 Current Capacity 402
    14.2.9.1 Wire Current Ratings 403
14.3 Insulation 403
  14.3.1 Plastics and Dielectric Constant 403
  14.3.2 Wire Insulation Characteristics 405
14.4 Jackets 405
14.5 Plastics 405
  14.5.1 Vinyl 405
  14.5.2 Polyethylene 405
  14.5.3 Teflon® 406
  14.5.4 Polypropylene 406
14.6 Thermoset Compounds 406
  14.6.1 Silicone 406
  14.6.2 Neoprene 406
  14.6.3 Rubber 406
  14.6.4 EPDM 406
14.7 Single Conductor 406
14.8 Multiconductor 407
  14.8.1 Multiconductor Insulation Color Codes 407
14.9 Pairs and Balanced Lines 407
  14.9.1 Multipair 408
  14.9.2 Analog Multipair Snake Cable 408
  14.9.3 High Frequency Pairs 409
    14.9.3.1 DVI 409
    14.9.3.2 HDMI 410
    14.9.3.3 IEEE -1394 or FireWire Serial Digital 410

14.9.3.4 USB . . . 411
14.9.3.5 DisplayPort . . . 411
14.9.3.6 Premise/Data Category Cables . . . 412
14.9.3.6.1 Cabling Definitions . . . 412
14.9.3.6.2 Structured Cabling . . . 412
14.9.3.6.3 Types of Structured Cables . . . 412
14.9.3.6.4 Comparisons . . . 413
14.9.3.6.5 Critical Parameters . . . 414
14.9.3.6.6 Terminating Connectors . . . 416
14.9.3.6.7 Baluns . . . 416
14.9.3.6.8 Adaptors . . . 417
14.9.3.6.9 Power Over Ethernet (PoE) . . . 417
14.9.3.6.10 Power Over Ethernet Plus (PoE Plus) . . . 417
14.10 Coaxial Cable . . . 418
14.10.1 History of Coaxial Cable . . . 418
14.10.2 Coaxial Cable Construction . . . 418
14.10.2.1 CCTV Cable . . . 419
14.10.2.1.1 CCTV Distances . . . 419
14.10.2.2 CATV Broadband Cable . . . 419
14.10.3 Coaxial Cable Installation Considerations . . . 419
14.10.3.1 Indoor Installation . . . 419
14.10.3.2 Outdoor Installation . . . 420
14.10.4 Coaxial Cable Termination Techniques . . . 420
14.10.4.1 Soldering . . . 420
14.10.4.2 Crimping . . . 420
14.10.4.3 Twist-On Connectors . . . 420
14.10.4.4 Compression Connectors . . . 420
14.10.5 Return Loss . . . 420
14.10.6 Video Triaxial Cable . . . 421
14.10.7 S-Video . . . 421
14.10.8 RGB . . . 421
14.10.9 VGA and Family . . . 422
14.11 Digital Video . . . 422
14.11.1 Digital Signals and Digital Cable . . . 422
14.11.2 Coax and SDI . . . 422
14.11.3 Cables and SDI . . . 422
14.11.4 Receiver Quality . . . 423
14.11.5 Serial Digital Video . . . 423
14.12 Radio Guide Designations . . . 424
14.13 Velocity of Propagation . . . 424
14.14 Shielding . . . 425
14.14.1 Serve or Spiral Shields . . . 425
14.14.2 Double Serve Shields . . . 425
14.14.3 French BraidTM . . . 426
14.14.4 Braid . . . 426
14.14.5 Foil . . . 426
14.14.6 Combination Shields . . . 426
14.14.6.1 Foil + Serve . . . 427
14.14.6.2 Foil + Braid . . . 427
14.14.6.3 Foil + Braid + Foil . . . 427

14.14.6.4 Foil + Braid + Foil + Braid 427
14.15 Shield Current Induced Noise 427
14.16 Grounds of Shields 427
14.16.1 Ground Loops 428
14.16.2 Telescopic Grounds 428
14.17 UTP and Audio 429
14.17.1 So Why Shields? 429
14.18 AES/EBU Digital Audio Cable 430
14.18.1 AES/EBU Digital Coaxial Cable 430
14.19 Triboelectric Noise 430
14.20 Conduit Fill 431
14.21 Long Line Audio Twisted Pairs 431
14.22 Delay and Delay Skew 431
14.23 Attenuation 432
14.24 Characteristic Impedance 432
14.25 Characteristic Impedance 433
14.26 Twisted-Pair Impedance 434
14.26.1 Transmission Line Termination 435
14.27 Loudspeaker Cable 435
14.27.1 Damping Factor 437
14.27.2 Crosstalk 437
14.28 National Electrical Code 438
14.28.1 Designation and Environmental Areas 439
14.28.2 Cable Types 440
14.28.3 NEC Substitution Chart 440
14.28.4 Final Considerations 440
14.28.5 Plenum Cable 441
14.28.6 Power Distribution Safety 442
14.28.6.1 Ground-Fault Interrupters 442
14.28.7 AC Power Cords and Receptacles 443
14.28.8 Shielded Power Cable 444
14.28.9 International Power 445
Acknowledgments 448
Bibliography 448

This page intentionally left blank

14.1 Introduction

It was not long ago that wire was the only method to inexpensively and reliably transmit sound or pictures from one place to another. Today we not only have wire, but we also have fiber optics, and wireless radio frequency (RF) transmission from Blu-tooth to wireless routers, cell phones, and microwave and satellite delivery. RF transmission is discussed briefly in Chapter 16.10, Wireless Microphones. This chapter will discuss the various forms of wire and cable used in audio and video.

Wire is a single conductive element. Wire can be insulated or uninsulated. Cable, on the other hand, is two or more conductive elements. While they theoretically could be uninsulated, the chance of them touching each other and creating a short circuit requires that they are usually both insulated. A cable can be multiple insulated wires, called a *multiconductor cable*, or wires that are twisted together, called a *twisted pair cable*, or cables with one wire in the center, surrounded by insulation and then a covering of metal used as another signal path, called *coaxial cable*.

14.2 Conductors

Wire and cable are used to connect one circuit or component to another. They can be internal, connecting one circuit to another inside a box, or externally connecting one box to another.

14.2.1 Resistance and Wire Size

Wire is made of metal, or other conductive compounds. All wire has resistance which dissipates power through heat. While this is not apparent on cables with small signals, such as audio or video signals, it is very apparent where high power or high current travels down a cable, such as a power cord. Resistance is related to the size of the wire. The smaller the wire, the greater the resistance.

14.2.2 Calculating Wire Resistance

The resistance for a given length of wire is determined by:

$$R = \frac{KL}{d^2} \tag{14-1}$$

where,

R is the resistance of the length of wire in ohms,

K is the resistance of the material in ohms per circular mil foot,

L is the length of the wire in feet,

d is the diameter of the wire in mils.

The resistance, in ohms per circular mil foot (Ω/cir mil ft), of many of the materials used for conductors is given in Table 14-1. The resistance shown is at 20°C (68°F), commonly called *room temperature*.

Table 14-1. Resistance of Metals and Alloys

Material	Symbol	Resistance (Ω/cir mil ft)
Silver	Ag	9.71
Copper	Cu	10.37
Gold	Au	14.55
Chromium	Cr	15.87
Aluminum	Al	16.06
Tungsten	W	33.22
Molybdenum	Mo	34.27
High-brass	Cu-Zn	50.00
Phosphor-bronze	Sn-P-Cu	57.38
Nickel, pure	Ni	60.00
Iron	Fe	60.14
Platinum	Pt	63.80
Palladium	Pd	65.90
Tin	Sn	69.50
Tantalum	Ta	79.90
Manganese-nickel	Ni-Mn	85.00
Steel	C-Fe	103.00
Lead	Pb	134.00
Nickel-silver	Cu-Zn-Ni	171.00
Alumel	Ni-Al-Mn-Si	203.00
Arsenic	As	214.00
Monel	Ni-Cu-Fe-Mn	256.00
Manganin	Cu-Mn-Ni	268.00
Constantan	Cu-Ni	270.00
Titanium	Ti	292.00
Chromel	Ni-Cr	427.00
Steel, manganese	Mn-C-Fe	427.00
Steel, stainless	C-Cr-Ni-Fe	549.00
Chromax	Cr-Ni-Fe	610.00
Nichrome V	Ni-Cr	650.00
Tophet A	Ni-Cr	659.00
Nichrome	Ni-Fe-Cr	675.00
Kovar A	Ni-Co-Mn-Fe	1732.00

When determining the resistance of a twisted pair, remember that the length of wire in a pair is twice the length of a single wire. Resistance in other constructions, such as coaxial cables, can be difficult to determine from just knowing the constituent parts. The center conductor might be easy to determine but a braid or braid + foil shield can be difficult. In those cases, consult the manufacturer.

Table 14-1 show the resistance in ohms (Ω) per foot per circular mil area for various metals, and combinations of metals (alloys). Of the common metals, silver is the lowest resistance. But silver is expensive and hard to work with. The next material, copper, is significantly less expensive, readily available, and lends itself to being annealed, which is discussed in Section 14.2.4. Copper is therefore the most common material used in the manufacture of wire and cable. However, where price is paramount and performance not as critical, aluminum is often used. The use of aluminum as the conducting element in a cable should be an indication to the user that this cable is intended to be lower cost and possibly lower performance.

One exception to this rule might be the use of aluminum foil which is often used in the foil shielding of even expensive high-performance cables. Another exception is emerging for automobile design, where the weight of the cable is a major factor. Aluminum is significantly less weight than copper, and the short distances required in cars means that resistance is less of a factor.

Table 14-1 may surprise many who believe, in error, that gold is the best conductor. The advantage of gold is its inability to oxidize. This makes it an ideal covering for articles that are exposed to the atmosphere, pollution, or moisture such as the pins in connectors or the connection points on insertable circuit boards. As a conductor, gold does not require annealing, and is often used in integrated circuits since it can be made into very fine wire. But, in normal applications, gold would make a poor conductive material, closer to aluminum in performance than copper.

One other material on the list commonly found in cable is steel. As can be seen, this material is almost ten times the resistance of copper, so many are puzzled by its use. In fact, in the cables that use steel wires, they are coated with a layer of copper, called copper-clad steel and signal passes only on the copper layer, an effect called skin effect that will be discussed in Section 14.2.8. Therefore, the steel wire is used for strength and is not intended to carry signals.

Copper-clad steel is also found in cables where cable pulling strength (pulling tension) is paramount. Then a stranded conductor can be made up of many copper-clad steel strands to maximize strength. Such a cable would compromise basic resistive performance. As is often the case, one can trade a specific attribute for another. In this case, better strength at the cost of higher resistance.

14.2.3 Resistance and Gage Size

In the United States, wire is sized by the American Wire Gage (AWG) method. AWG was based on the previous Brown and Sharpe (B & S) system of wire sizes which dates from 1856. AWG numbers are most common in the United States, and will be referred to throughout this book. The wire most often used in audio ranges from approximately 10 AWG to 30 AWG, although larger and smaller gage sizes exist. Wire with a small AWG number, such as 4 AWG, is very heavy, physically strong but cumbersome, and has very low resistance, while wire of larger numbers, such as 30 AWG can be very light weight and fragile, and has high resistance. Resistance is an important factor in determining the appropriate wire size in any circuit. For instance, if an 8 Ω loudspeaker is being connected to an amplifier 500 ft away through a #19 wire, 50% of the power would be dropped in the wire in the form of heat. This is discussed in Section 14.25 regarding loudpeaker cable.

Each time the wire size changes three numbers, such as from 16 AWG to 19 AWG the resistance doubles. The reverse is also true. With a wire changed from 16 AWG to 13 AWG, the resistance halves. This also means that combining two identical wires of any given gage decreases the total gage of the combined wires by three units, and reduces the resistance. Two 24 AWG wires combined (twisted together) would be 21 AWG, for instance. If wires are combined of different gages, the resulting gage can be easily calculated by adding the circular mil area (CMA) shown in Tables 14-2 and 14-3. For instance, if three wires were combined, one 16 AWG (2583 CMA), one 20 AWG (1022 CMA) and one 24 AWG (404 CMA), the total CMA would be 2583 + 1022 + 404 = 4009 CMA. Looking in Table 14-1, this numbers falls just under 14 AWG. While even number gages are the most common, odd number gages (e.g., 23 AWG) can sometimes be found. There are many Category 6 (Cat 6) premise/data cables that are 23 AWG, for instance. When required, manufacturers can even produce partial gages. There are coaxial cables with 28.5 AWG center conductors. Such specialized gage sizes might require equally special connectors.

There are two basic forms of wire, solid and stranded. A solid conductor is one continuous piece of metal. A stranded conductor is made of multiple smaller wires combined to make a single conductor. Solid wire has slightly lower resistance, with less flexibility and less flex-life (flexes to failure) than stranded wire.

14.2.4 Drawing and Annealing

Copper conductors start life as copper ore in the ground. This ore is mined, refined, and made into bars or rod. Five sixteenth inch copper rod is the most common form used for the making of wire and cable. Copper can be purchased at various purities. These commonly follow the ASTM (American Society for Testing and Materials) standards. Most of the high-purity copper is known as ETP, electrolytic tough pitch. For example, many cable products are manufactured with ASTM B115 ETP. This copper is 99.95% pure. Copper of higher purity can be purchased should the requirement arise. Many consumer audiophiles consider these to be oxygen free, when this term is really a discussion of copper purity and is determined by the number of nines of purity. The cost of the copper rises dramatically with each “9” that is added.

To turn $^5/_{16}$ inch rod into usable wire, the copper rod is drawn through a series of dies. Each time it makes the rod slightly smaller. Eventually you can work the rod down to a very long length of very small wire. To take $^5/_{16}$ inch rod down to a 12 AWG wire requires drawing the conductor through eleven different dies. Down to 20 AWG requires fifteen dies. To take that wire down to 36 AWG requires twenty-eight dies.

The act of drawing the copper work hardens the material making it brittle. The wire is run through an in-line annealing oven, at speeds up to 7000 feet per minute, and a temperature of 900 to 1000°F (482 to 537°C). This temperature is not enough to melt the wire, but it is enough to let the copper lose its brittleness and become flexible again, to reverse the work hardening. Annealing is commonly done at the end of the drawing process. However, if the next step requires more flexibility, it can be annealed partway through the drawing process. Some manufacturers draw down the wire and then put the entire roll in an annealing oven. In order to reduce oxygen content, some annealing ovens have inert atmospheres, such as nitrogen. This increases the purity of the copper by reducing the oxygen content. But in-line annealing is more consistent than a whole roll in an oven.

Lack of annealing, or insufficient annealing time or temperature, can produce a conductor which is stiff, brittle, and prone to failure. With batch annealing, the inner windings in a roll may not be heated as effectively as the outer windings. Cables made in other countries may not have sufficient purity for high-performance applications. Poor--quality copper, or poor annealing, are very hard to tell from initial visual inspection but often shows up during or after installation.

14.2.5 Plating and Tinning

Much of the wire manufactured is plated with a layer of tin. This can also be done in-line with the drawing and annealing by electroplating a layer on the wire. Tinning makes the wire especially resistant to pollutants, chemicals, salt (as in marine applications). But such a plated conductor is not appropriate for high-frequency applications where the signal travels on the skin of the conductor, called *skin effect*. In that case, bare copper conductors are used. The surface of a conductor used for high frequencies is a major factor in good performance and should have a mirror finish on that surface. Wires are occasionally plated with silver. While silver is slightly more conductive, its real advantage is that silver oxide is the same resistance as bare silver. This is not true with copper, where copper oxide is a semiconductor. Therefore, where reactions with a copper wire are predicted, silver plating may help preserve performance. So silver plating is sometimes used for marine cables, or cables used in similar outdoor environments.

Some plastics, when extruded (melted) onto wires, can chemically affect the copper. This is common, for instance, with an insulation of extruded TFE (tetrafluoroethylene), a form of Teflon™. Wires used inside these cables are often silver plated or silver-clad. Any oxidizing caused by the extrusion process therefore has no effect on performance. Of course, just the cost of silver alone makes any silver-plated conductor significantly more expensive than bare copper.

14.2.6 Conductor Parameters

Table 14-2 shows various parameters for solid wire from 4 AWG to 40 AWG. Table 14-3 shows the same parameters for stranded wire. Note that the resistance of a specific gage of solid wire is lower than stranded wire of the same gage. This is because the stranded wire is not completely conductive; there are spaces (interstices) between the strands. It takes a larger stranded wire to equal the resistance of a solid wire.

Table 14-2. Parameters for Solid Wire from 4 AWG to 40 AWG

AWG	Nominal Diameter	CMA (×1000)	Bare lbs/ft	Ω/100 0 ft	Current A	MM^2 Equivalent
4	0.2043	41.7	0.12636	0.25	59.57	21.1
5	0.1819	33.1	0.10020	0.31	47.29	16.8
6	0.162	26.3	0.07949	0.4	37.57	13.3
7	0.1443	20.8	0.06301	0.5	29.71	10.6
8	0.1285	16.5	0.04998	0.63	23.57	8.37
9	0.1144	13.1	0.03964	0.8	18.71	6.63
10	0.1019	10.4	0.03143	1	14.86	5.26
11	0.0907	8.23	0.02493	1.26	11.76	4.17
12	0.0808	6.53	0.01977	1.6	9.33	3.31
13	0.075	5.18	0.01567	2.01	7.40	2.62
14	0.0641	4.11	0.01243	2.54	5.87	2.08
15	0.0571	3.26	0.00986	3.2	4.66	1.65
16	0.0508	2.58	0.00782	4.03	3.69	1.31
17	0.0453	2.05	0.00620	5.1	2.93	1.04
18	0.0403	1.62	0.00492	6.4	2.31	0.823
19	0.0359	1.29	0.00390	8.1	1.84	0.653
20	0.032	1.02	0.00309	10.1	1.46	0.519
21	0.0285	0.81	0.00245	12.8	1.16	0.412
22	0.0254	0.642	0.00195	16.2	0.92	0.324
23	0.0226	0.51	0.00154	20.3	0.73	0.259
24	0.0201	0.404	0.00122	25.7	0.58	0.205
25	0.0179	0.32	0.00097	32.4	0.46	0.162
26	0.0159	0.253	0.00077	41	0.36	0.128
27	0.0142	0.202	0.00061	51.4	0.29	0.102
28	0.0126	0.159	0.00048	65.3	0.23	0.08
29	0.0113	0.127	0.00038	81.2	0.18	0.0643
30	0.01	0.1	0.00030	104	0.14	0.0507
31	0.0089	0.0797	0.00024	131	0.11	0.0401
32	0.008	0.064	0.00019	162	0.09	0.0324
33	0.0071	0.0504	0.00015	206	0.07	0.0255
34	0.0063	0.0398	0.00012	261	0.06	0.0201
35	0.0056	0.0315	0.00010	331	0.05	0.0159
36	0.005	0.025	0.00008	415	0.04	0.0127
37	0.0045	0.0203	0.00006	512	0.03	0.0103
38	0.004	0.016	0.00005	648	0.02	0.0081
39	0.0035	0.0123	0.00004	847	0.02	0.0062
40	0.003	0.0096	0.00003	1080	0.01	0.0049

14.2.6.1 Stranded Cables.

Stranded cables are more flexible, and have greater flex-life (flexes to failure) than solid wire. Table 14-4 shows some suggested construction values. The two numbers (65 × 34, for example) show the number of strands (65) and the gage size of each strand (34) for each variation in flexing.

Table 14-3. Parameters for ASTM Class B Stranded Wires from 4 AWG to 40 AWG

AWG	Nominal Diameter	CMA (×1000)	Bare lbs/ft	Ω/ 1000 ft	Current A*	MM2 Equivalent
4	0.232	53.824	0.12936	0.253	59.63	27.273
5	0.206	42.436	0.10320	0.323	47.27	21.503
6	0.184	33.856	0.08249	0.408	37.49	17.155
7	0.164	26.896	0.06601	0.514	29.75	13.628
8	0.146	21.316	0.05298	0.648	23.59	10.801
9	0.13	16.9	0.04264	0.816	18.70	8.563
10	0.116	13.456	0.03316	1.03	14.83	6.818
11	0.103	10.609	0.02867	1.297	11.75	5.376
12	0.0915	8.372	0.02085	1.635	9.33	4.242
13	0.0816	6.659	0.01808	2.063	8.04	3.374
14	0.0727	5.285	0.01313	2.73	5.87	2.678
15	0.0647	4.186	0.01139	3.29	4.66	2.121
16	0.0576	3.318	0.00824	4.35	3.69	1.681
17	0.0513	2.632	0.00713	5.25	2.93	1.334
18	0.0456	2.079	0.00518	6.92	2.32	1.053
19	0.0407	1.656	0.00484	8.25	1.84	0.839
20	0.0362	1.31	0.00326	10.9	1.46	0.664
21	0.0323	1.043	0.00284	13.19	1.16	0.528
22	0.0287	0.824	0.00204	17.5	0.92	0.418
23	0.0256	0.655	0.00176	20.99	0.73	0.332
24	0.0228	0.52	0.00129	27.7	0.58	0.263
25	0.0203	0.412	0.01125	33.01	0.46	0.209
26	0.018	0.324	0.00081	44.4	0.36	0.164
27	0.0161	0.259	0.00064	55.6	0.29	0.131
28	0.0143	0.204	0.00051	70.7	0.23	0.103
29	0.0128	0.164	0.00045	83.99	0.18	0.083
30	0.0113	0.128	0.00032	112	0.14	0.0649
31	0.011	0.121	0.00020	136.1	0.11	0.0613
32	0.009	0.081	0.00020	164.1	0.09	0.041
33	0.00825	0.068	0.00017	219.17	0.07	0.0345
34	0.0075	0.056	0.00013	260.9	0.06	0.0284
35	0.00675	0.046	0.00011	335.96	0.04	0.0233
36	0.006	0.036	0.00008	414.8	0.04	0.0182
37	0.00525	0.028	0.00006	578.7	0.03	0.0142
38	0.0045	0.02	0.00005	658.5	0.02	0.0101
39	0.00375	0.014	0.00004	876.7	0.02	0.0071
40	0.003	0.009	0.00003	1028.8	0.01	0.0046

*For both solid and stranded wire, amperage is calculated at 1 A for each 700 CMA. See also Section 14.2.9.

Table 14-4. Suggested Conductor Strandings for Various Degrees of Flexing Severity

Typical Applications	AWG	mm	AWG	mm
	12 AWG		**14 AWG**	
Fixed Service (Hook-Up Wire Cable in Raceway)			Solid	
	19 × 25	19 × 0.455	19 × 27	19 × 0.361
Moderate Flexing (Frequency Disturbed for Maintenance)	65 × 30	65 × 0.254	19 × 27	19 × 0.361
			41 × 30	41 × 0.254
Severe Flexing (Microphones and Test Prods)	165 × 34	165 × 0.160	104 × 34	104 × 0.160
	16 AWG		**18 AWG**	
Fixed Service (Hook-Up Wire Cable in Raceway)	Solid		Solid	
	19 × 29	19 × 0.287	7 × 26	7 × 0.404
			16 × 30	16 × 0.254
Moderate Flexing (Frequently Disturbed for Maintenance)	19 × 29	19 × 0.287	16 × 30	16 × 0.254
	26 × 30	26 × 0.254	41 × 34	41 × 0.160
Severe Flexing (Microphones, Test Prods)	65 × 34	65 × 0.160	41 × 34	41 × 0.160
	104 × 36	104 × 0.127	65 × 36	65 × 0.127
	20 AWG		**22 AWG**	
Fixed Service (Hook-Up Wire Cable in Raceway)	Solid		Solid	
	7 × 28	7 × 0.320	7 × 30	7 × 0.254
	10 × 30	10 × 0.254		
Moderate Flexing (Frequency Distributed for Maintenance)	7 × 28	7 × 0.320	7 × 30	7 × 0.254
	10 × 30	10 × 0.254		
	19 × 32	19 × 0.203	19 × 34	19 × 0.160
	26 × 34	26 × 0.160		
Severe Flexing (Microphones, Test Prods)	26 × 34	26 × 0.160	19 × 34	19 × 0.160
	42 × 36	42 × 0.127	26 × 36	26 × 0.127
	24 AWG		**26 AWG**	
Fixed Service (Hook-Up Wire Cable in Raceway)	Solid		Solid	
	7 × 32	7 × 0.203	7 × 34	7 × 0.160
Moderate Flexing (Frequently Disturbed for Maintenance)	7 × 32	7 × 0.203	7 × 34	7 × 0.160
	10 × 34	10 × 0.160		
Severe Flexing (Microphones, Test Prods)	19 × 36	19 × 0.127	7 × 34	7 × 0.160
	45 × 40	45 × 0.079	10 × 36	10 × 0.127

Courtesy Belden

14.2.7 Pulling Tension

Pulling tension must be adhered to so the cable will not be permanently elongated. The pulling tension for annealed copper conductors is shown in Table 14-5.

Multiconductor cable pulling tension can be determined by multiplying the total number of conductors by the appropriate value. For twisted pair cables, there are two wires per pair. For shielded twisted pair cables, with foil shields, there is a drain wire that must be included in the calculations. Be cautious: the drain wire can sometimes be smaller gage than the conductors in the pair. The pulling tension of coaxial cables or other cables that are not multiple conductors is much harder to calculate. Consult the manufacturer for the required pulling tension.

Table 14-5. Pulling Tension for Annealed Copper Conductors

24 AWG	5.0 lbs
22 AWG	7.5 lbs
20 AWG	12.0 lbs
18 AWG	19.5 lbs
16 AWG	31.0 lbs
14 AWG	49.0 lbs
12 AWG	79.0 lbs

14.2.8 Skin Effect

As the frequency of the signal increases on a wire, the signal travels closer to the surface of the conductor. Since very little of the area of the center conductor is used at high frequencies, some cable is made with a copper clad-steel core center conductor. These are known as copper-clad, copper-covered, or Copperweld™ and are usually used by CATV/broadband service providers.

Copper-clad steel is stronger than copper cable so it can more easily withstand pulling during installation, or wind, ice, and other outside elements after installation. For instance, a copper-clad #18 AWG coaxial cable has a pull strength of 102 lbs while a solid copper #18 AWG coax would have a pull strength of 69 lbs. The main disadvantage is that steel is not a good conductor below 50 MHz, between four and seven times the resistance of copper, depending on the thickness of the copper layer.

This is a problem where signals are below 50 MHz such as DOCSIS data delivery, or VOD (video-on-demand) signals which are coming from the home to the provider. When installing cable in a system, it is better to use solid copper cable so it can be used at low frequencies as well as high frequencies.

This is also why copper-clad conductors are not appropriate for any application below 50 MHz, such as baseband video, CCTV, analog, or digital audio. Copper-clad is also not appropriate for applications such as SDI or HD-SDI video, and similar signals where a significant portion of the data is below 50 MHz.

The skin depth for copper conductors can be calculated with the equation

$$D = \frac{2.61}{\sqrt{f}} \tag{14-2}$$

where,
D is the skin depth in inches,
f is the frequency in hertz.

Table 14-6 compares the actual skin depth and percent of the center conductor actually used in an RG-6 cable. The skin depth always remains the same no matter what the thickness of the wire is. The only thing that changes is the percent of the conductor utilized. Determining the percent of the conductor utilized requires using two times the skin depth because we are comparing the diameter of the conductor to its depth.

As can be seen, by the time the frequencies are high, the depth of the signal on the skin can easily be micro-inches. For signals in that range, such as high-definition video signals, for example, this means that the surface of the wire is as critical as the wire itself. Therefore, conductors intended to carry high frequencies should have a mirror finish.

Table 14-6. Skin Depths at Various Frequencies

Frequency	Skin Depth in Inches	% Used of #18 AWG Conductor
1 kHz	0.082500	100
10 kHz	0.026100	100
100 kHz	0.008280	41
1 MHz	0.002610	13
10 MHz	0.000825	4.1
100 MHz	0.000261	1.3
1 GHz	0.0000825	0.41
10 GHz	0.0000261	0.13

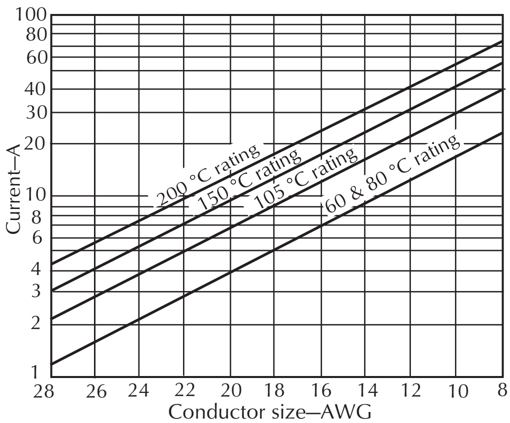
Since the resistance of the wire at these high frequencies is of no consequence, it is sometimes asked why larger conductors go farther. The reason is that the *surface area*, the *skin*, on a wire is greater as the wire gets larger in size.

Further, some conductors have a tin layer to help prevent corrosion. These cables are obviously not intended for use at frequencies above just a few megahertz, or a significant portion of the signal would be traveling in the tin layer. Tin is not an especially good conductor as can be seen in Table 14-1.

14.2.9 Current Capacity

For conductors that will carry large amounts of electrical flow, large amperage or current from point to point, a general chart has been made to simplify the current carrying capacity of each conductor. To use the current capacity chart in Fig. 14-1, first determine conductor gage, insulation and jacket temperature rating, and number of conductors from the applicable product description for the cable of interest. These can usually be obtained from a manufacturer's Web site or catalog.

Next, find the current value on the chart for the proper temperature rating and conductor size. To calculate the maximum current rating/conductor multiply the chart value by the appropriate conductor factor. The chart assumes the cable is surrounded by still air at an ambient temperature of 25°C (77°F). Current values are in amperes (rms) and are valid for copper conductors only. The maximum continuous current rating for an electronic cable is limited by conductor size, number of conductors contained within the cable, maximum temperature rating of the insulation on the conductors,



# Conductors*	Factor	# Conductors*	Factor
1	1.6	6–15	0.7
2–3	1.0	16–30	0.5
4–5	0.8		

* do not count shields unless used as a conductor

Figure 14-1. Current ratings for electronic cable. Courtesy Belden.

and environment conditions such as ambient temperature and air flow. The four lines marked with temperatures apply to different insulation plastics and their melting point. Consult the manufacturer's Web site or catalog for the maximum insulation or jacket temperature.

The current ratings of Fig. 14-1 are intended as general guidelines for low-power electronic communications and control applications. Current ratings for high-power applications generally are set by regulatory agencies such as Underwriters Laboratories (UL), Canadian Standards Association (CSA), National Electrical Code (NEC), and others and should be used before final installation.

Table 310-15(b)(2)(a) of the NEC contains amperage adjustment factors for whenever more than three current carrying conductors are in a conduit or raceway.

Section 240-3 of the NEC provides requirements for overload protection for conductors other than flexible cords and fixture wires. Section 240-3(d), Small Conductors, states that #14 to #10 conductors require a maximum protective overcurrent device with a rating no higher than the current rating listed in the 60°C column. These currents are 15 A for #14 copper wire, 20 A for #12 copper wire, and 30 A for #10 copper wire. These values are familiar as the breaker ratings for commercial installations.

When connecting wire to a terminal strip or another wire etc., the temperature rise in the connections must also be taken into account. Often the circuit is not limited by the current carrying capacity of the wire but of the termination point.

14.2.9.1 Wire Current Ratings

Current carrying capacity of wire is controlled by the NEC, particularly in Table 310-16, Table 310-15(b)(2)(a), and Section 240-3.

Table 310-16 of the NEC shows the maximum current carrying capacity for insulated conductors rated from 0 to 2000 V, including copper and aluminum conductors. Each conductor amperage is given for three temperatures: 60°C, 75°C, and 90°C. Copper doesn't melt until almost 2000°, so the current limit on a copper wire is not the melting point of the wire but the melting point of the insulation. This number is listed by most manufacturers in their catalog or on their Web site. For instance, PVC (polyvinyl chloride) can be formulated to withstand temperatures from 60°C to as high as 105°C. The materials won't melt right at the specified temperature, but may begin to fail certain tests, such as cracking when bent.

14.3 Insulation

Wire can be bare, often called *bus bar* or *bus wire*, but is most often insulated. It is covered with a non-conducting material. Early insulations included cotton or silk woven around the conductor, or even paper. Cotton-covered house wiring can still be found in perfect operating condition in old houses. Today, most insulation materials are either some kind of rubber or some kind of plastic. The material chosen should be listed in the manufacturer's catalog with each cable type. Table 14-7 lists some of the rubber-based materials with their properties. Table 14-8 lists the properties of various plastics. The ratings in both tables are based on average performance of general-purpose compounds. Any given property can usually be improved by the use of selective compounding.

14.3.1 Plastics and Dielectric Constant

Table 14-9 is a list of various insulation materials with details on performance, requirements, and special advantages. Insulation, when used on a cable intended to carry a signal, is often referred to as a *dielectric*. The performance of any material, its ability to insulate with minimal effect to the signal running on the cable is called the *dielectric constant* and can be measured in a

Table 14-7. Comparative Properties of Rubber Insulation. (Courtesy Belden)

Properties	Rubber	Neoprene	Hypalon (Chlorosulfonated Polyethylene)	EPDM (Ethylene Propylene Diene Monomer)	Silicone
Oxidation Resistance	F	G	E	G	E
Heat Resistance	F	G	E	E	O
Oil Resistance	P	G	G	F	F-G
Low Temperature Flexibility	G	F-G	F	G-E	O
Weather, Sun Resistance	F	G	E	E	O
Ozone Resistance	P	G	E	E	O
Abrasion Resistance	E	G-E	G	G	P
Electrical Properties	E	P	G	E	O
Flame Resistance	P	G	G	P	F-G
Nuclear Radiation Resistance	F	F-G	G	G	E
Water Resistance	G	E	G-E	G-E	G-E
Acid Resistance	F-G	G	E	G-E	F-G
Alkali Resistance	F-G	G	E	G-E	F-G
Gasoline, Kerosene, etc. (Aliphatic Hydrocarbons) Resistance	P	G	F	P	P-F
Benzol, Toluol, etc.(Aromatic Hydrocarbons) Resistance	P	P-F	F	F	P
Degreaser Solvents (Halogenated Hydrocarbons) Resistance	P	P	P-F	P	P-G
Alcohol Resistance	G	F	G	P	G

P = poor, F = fair, G = good, E = excellent, O = outstanding

Table 14-8. Comparative Properties of Plastic Insulation. (Courtesy Belden)

Properties	PVC	Low-Density Polyethylene	Cellular Polyethylene	High-Density Polyethylene	Polyethylene	Polyurethane	Nylon	Teflon®
Oxidation Resistance	E	E	E	E	E	E	E	O
Heat Resistance	G-E	G	G	E	E	G	E	O
Oil Resistance	F	G	G	G-E	F	E	E	O
Low Temperature Flexibility	P-G	G-E	E	E	P	G	G	O
Weather, Sun Resistance	G-E	E	E	E	E	G	E	O
Ozone Resistance	E	E	E	E	E	E	E	E
Abrasion Resistance	F-G	F-G	F	E	F-G	O	E	E
Electrical Properties	F-G	E	E	E	E	P	P	E
Flame Resistance	E	P	P	P	P	P	P	O
Nuclear Radiation Resistance	G	G	G	G	F	G	F-G	P
Water Resistance	E	E	E	E	E	P-G	P-F	E
Acid Resistance	G-E	G-E	G-E	G-E	E	F	P-F	E
Alkali Resistance	G-E	G-E	G-E	G-E	E	F	E	E
Gasoline, Kerosene, etc. (Aliphatic Hydrocarbons) Resistance	P	P-F	P-F	P-F	P-F	G	G	E
Benzol, Toluol, etc. (Aromatic Hydrocarbons) Resistance	P-F	P	P	P	P-F	P	G	E
Degreaser Solvents (Halogenated Hydrocarbons) Resistance	P-F	P	P	P	P	P	G	E
Alcohol Resistance	G-E	E	E	E	E	P	P	E

P = poor, F = fair, G = good, E = excellent, O = outstanding laboratory. Table 14-9 shows some standard numbers as a point of reference.

Table 14-9. Dielectric Constant

Dielectric Constant	Material	Note
1	Vacuum	By definition
1.0167	Air	Very close to 1
1.35	Foam, Air-filled Plastic	Current technological limit
2.1	Solid Teflon™	Best solid plastic
2.3	Solid Polyethylene	Most common plastic
3.5–6.5	Solid Polyvinyl Chloride	Low price, easy to work with

14.3.2 Wire Insulation Characteristics

The key difference between rubber compounds and plastic compounds is their recyclability. Plastic materials can be ground up, and re-melted into other objects. Polyethylene, for instance, can be recycled into plastic bottles, grocery bags, or even park benches. And, should the need arise, these objects could themselves be ground up and turned back into wire insulation, or many other uses. The term *thermoplastic* means changed by heat and is the source of the common term *plastic*.

Rubber compounds, on the other hand, are thermoset. That is, once they are made, they are set, and the process cannot be reversed. Rubber, and its family, is cured in a process sometimes called *vulcanizing*. These compounds cannot be ground up and recycled into new products. There are natural rubber compounds (such as latex-based rubber) and artificial, chemical-based rubber compounds such as EPDM (ethylene-propylene-diene monomer).

The vast majority of wire and cable insulations are plastic-based compounds. Rubber, while it is extremely rugged, is considerably more expensive that most plastics, so there are fewer and fewer manufacturers offering rubber-based products. These materials, both rubber and plastic, are used in two applications with cable. The first application is insulation of the conductor(s) inside the cable. The second is as a jacket material to protect the contents of the cable.

14.4 Jackets

The jacket characteristics of cable have a large effect on its ruggedness and the effect of environment. A key consideration is often flexibility, especially at low temperatures. Audio and broadcast cables are manufactured in a wide selection of standard jacketing materials. Special compounds and variations of standard compounds are used to meet critical audio and broadcast application requirements and unusual environmental conditions. Proper matching of cable jackets to their working environment can prevent deterioration due to intense heat and cold, sunlight, mechanical abuse, impact, and crowd or vehicle traffic.

14.5 Plastics

Plastic is a shortened version of the term *thermoplastic*. *Thermo* means heat, *plastic* means change. Thermoplastic materials can be changed by heat. They can be melted and extruded into other shapes. They can be extruded around wires, for instance, forming an insulative (nonconductive) layer. There are many forms of plastic. Below is a list of the most common varieties used in the manufacture of wire and cable.

14.5.1 Vinyl

Vinyl is sometimes referred to as PVC or *polyvinyl chloride*, and is a chemical compound invented in 1928 by Dr. Waldo Semon (USA). Extremely high or low temperature properties cannot be found in one formulation, therefore, formulations may have −55°C to +105°C (−67°F to +221°F) rating while other common vinyls may have −20°C to +60°C (−4°F to +140°F). The many varieties of vinyl also differ in pliability and electrical properties fitting a multitude of applications. The price range can vary accordingly. Typical dielectric constant values can vary from 3.5 at 1000 Hz to 6.5 at 60 Hz, making it a poor choice if high performance is required. PVC is one of the least expensive compounds, and one of the easiest to work with. Therefore, PVC is used with many cables that do not require high performance, or where cost of materials is a major factor. PVC is easy to color, and can be quite flexible, although it is not very rugged. In high-performance cables, PVC is often used as the jacket material, but not inside the cable.

14.5.2 Polyethylene

Polyethylene, invented by accident in 1933 by E.W. Fawcett and R.O. Gibson (Great Britain), is a very good insulation in terms of electrical properties. It has a low dielectric constant value over all frequencies and very high insulation resistance. In terms of flexibility, polyethylene can be rated stiff to very hard depending on molecular weight and density. Low density is the most flexible and high density high molecular weight formulations are very hard. Moisture resistance is rated excellent. Correct brown and black formulations have excellent sunlight resistance. The dielectric constant is 2.3 for solid insulation and as low as 1.35 for gas-injected foam cellular designs. Polyethylene is the most common plastic worldwide.

14.5.3 Teflon®

Invented in 1937 by Roy Plunkett (USA) at DuPont, Teflon has excellent electrical properties, temperature range, and chemical resistance. It is not suitable where subjected to nuclear radiation, and it does not have good high voltage characteristics. FEP (fluorinated ethylene-propylene) Teflon is extrudable in a manner similar to vinyl and polyethylene, therefore, long wire and cable lengths are available. TFE (tetrafluoroethylene) Teflon is extrudable in a hydraulic ram-type process and lengths are limited due to amount of material in the ram, thickness of the insulation, and core size. TFE must be extruded over silver-coated or nickel-coated wire. The nickel and silver-coated designs are rated +260°C and +200°C maximum (500°F and 392°F), respectively, which is the highest temperature for common plastics. The cost of Teflon is approximately eight to ten times more per pound than vinyl insulations. The dielectric constant for solid Teflon is 2.1, the lowest of all solid plastics. Foam Teflon (FEP) has a dielectric constant as low as 1.35. Teflon is produced by and a trademark of DuPont Corporation.

14.5.4 Polypropylene

Polypropylene is similar in electrical properties to polyethylene and is primarily used as an insulation material. Typically, it is harder than polyethylene, which makes it suitable for thin wall insulations. UL maximum temperature rating may be 60°C or 80°C (140°F or 176°F). The dielectric constant is 2.25 for solid and 1.55 for cellular designs.

14.6 Thermoset Compounds

As the name implies, *thermoset* compounds are produced by heat (thermo) but are set. That is, the process cannot be reversed as in thermoplastics. They cannot be recycled into new products as thermoplastic materials can.

14.6.1 Silicone

Silicone is a very soft insulation which has a temperature range from −80°C to +200°C (−112°F to +392°F). It has excellent electrical properties plus ozone resistance, low moisture absorption, weather resistance, and radiation resistance. It typically has low mechanical strength and poor scuff resistance. Silicone is seldom used because it is very expensive.

14.6.2 Neoprene

Neoprene has a maximum temperature range from −55°C to +90°C (−67°F to +194°F). The actual range depends on the formulation used. Neoprene is both oil and sunlight resistant making it ideal for many outdoor applications. The most stable colors are black, dark brown, and gray. The electrical properties are not as good as other insulation material; therefore, thicker insulation must be used for the same insulation.

14.6.3 Rubber

The description of *rubber* normally includes natural rubber and styrene-butadiene rubber (SBR) compounds. Both can be used for insulation and jackets. There are many formulations of these basic materials and each formulation is for a specific application. Some formulations are suitable for −55°C (−67°F) minimum while others are suitable for +75°C (+167°F) maximum. Rubber jacketing compounds feature exceptional durability for extended cable life. They withstand high-impact and abrasive conditions better than PVC and are resistant to degradation or penetration by water, alkali, or acid. They have excellent heat resistant properties, and also provide greater cable flexibility in cold temperatures.

14.6.4 EPDM

EPDM stands for *ethylene-propylene-diene monomer*. It was invented by Dr. Waldo Semon in 1927 (see Section 14.5.1). It is extremely rugged, like natural rubber, but can be created from petroleum byproducts ethylene and propylene gas.

14.7 Single Conductor

Single conductor wire starts with a single wire, either solid or stranded. It can be bare, sometimes called *buss bar*, or can be jacketed. There is no actual limit to how small, or how large, a conductor could be. Choice of size (AWG) will be based on application and the current or wattage delivery required. If jacketed, the choice of jacket can be based on performance, ruggedness, flexibility, or any other requirement.

There is no single conductor plenum rating because the NEC (National Electrical Code) only applies to cables, more than one conductor. However, Articles 300 and 310 of the NEC are sometimes cited when installing single conductor wire for grounds and similar applications.

14.8 Multiconductor

Bundles of two or more insulated wires are considered multiconductor cable. Besides the requirements for each conductor, there is often an overall jacket, chosen for whatever properties would be appropriate for a particular application.

There are specialized multiconductor cables, such as power cordage used to deliver ac power from a wall outlet (or other source) to a device. There are UL safety ratings on such a cable to assure users will not be harmed.

There are other multiconductor applications such as VFD (variable frequency drive) cables, specially formulated to minimize standing waves and arcing discharge when running variable frequency motors. Since a multiconductor cable is not divided into pairs, resistance is still the major parameter to be determined, although reactions between conductors (as in VFD) can also be considered.

14.8.1 Multiconductor Insulation Color Codes

The wire insulation colors help trace conductors or conductor pairs. There are many color tables; Table 14-10 is one example.

14.9 Pairs and Balanced Lines

Twisting two insulated wires together makes a twisted pair. Since two conductive paths are needed to make a circuit, twisted pairs give users an easy way to connect power or signals from point to point. Sometimes the insulation color is different to identify each wire in each pair. Pairs can have dramatically better performance than multiconductor cables because pairs can be driven as a balanced line.

A balanced line is a configuration where the two wires are electrically identical. The electrical performance is referred to ground, the zero point in circuit design. Balanced lines reject noise, from low frequencies, such as 50/60 Hz power line noise, up to radio frequency signals in the Megahertz, or even higher.

When the two conductors are electrically identical, or close to identical, there are many other parameters, besides resistance, that come into play. These include capacitance, inductance, and impedance. And when we get to high-frequency pairs, such as data cables, we even measure the variations in resistance (resistance unbalance), variations in capacitance (capacitance unbalance, or even variations in impedance (return loss). Each of these has a section farther on in this chapter.

Table 14-10. Color Code for Nonpaired Cables per ICEA #2 and #2R

Conductor	Color	Conductor	Color	Conductor	Color	Conductor	Color
1st	Black	14th	Green/White	27th	Blue/Blk/Wht	40th	Red/Wht/Grn
2nd	White	15th	Blue/White	28th	Blk/Red/Grn	41st	Grn/Wht/Blue
3rd	Red	16th	Black/Red	29th	Wht/Red/Grn	42nd	Org/Red.Grn
4th	Green	17th	White/Red	30th	Red/Blk/Grn	43rd	Blue/Red/Grn
5th	Orange	18th	Orange/Red	31st	Grn/Blk/Org	44th	Blk/Wht/Blue
6th	Blue	19th	Blue/Red	32nd	Org/Blk/Grn	45th	Wht/Blk/Blue
7th	White/Black	20th	Red/Green	33rd	Blue/Wht/Org	46th	Red/Wht/Blue
8th	Red/Black	21st	Orange/Green	34th	Blk/Wht/Org	47th	Grn/Orn/Red
9th	Green/Black	22nd	Blk/Wht/Red	35th	Wht/Red/Org	48th	Org/Red/Blue
10th	Orange/Black	23rd	Wht/Blk/Red	36th	Org/Wht/Blue	49th	Blue/Red/Org
11th	Blue/Black	24th	Red/Blk/Wht	37th	Wht/Red/Blue	50th	Blk/Org/Red
12th	Black/White	25th	Grn/Blk/Wht	38th	Blk/Wht/Grn		
13th	Red/White	26th	Org/Blk/Wht	39th	Wht/Blk/Grn		

Courtesy Belden

Balanced lines work because they have a transformer at each end, a device made of two coils of wire wound together. Many modern devices now use circuits that act electrically the same as a transformer, an effect called *active balancing*. The highest-quality transformers can be extremely expensive, so high-performing balanced-line chips have been improving, some getting very close to the coils-of-wire performance.

It should be noted that virtually all professional installations use twisted pairs for audio because of their noise rejection properties. In the consumer world, the cable has one hot connection and a grounded shield around it and is called an unbalanced cable. These cables are effective for only short distances and have no other inherent noise rejection besides the shield itself.

14.9.1 Multipair

As the name implies, multipair cables contain more than one pair. Sometimes referred to as *multicore* cables, these can just be grouped bare pairs, or each pair could be individually jacketed, or each pair could be shielded (shielding is outlined below), or the pairs could even be individually shielded and jacketed. All of these options are easily available. Where there is an overall jacket, or individual jackets for each pair, the jacket material for each pair is chosen with regard to price, flexibility, ruggedness, color, and any other parameter required.

It should be noted that the jackets on pairs, or the overall jacket, has almost no effect on the performance of the pairs. One could make a case that, with individually jacketed pairs, the jacket moves the pairs apart and therefore improves crosstalk between pairs. It is also possible that poorly extruded jackets could leak the chemicals that make up the jacket into the pair they are protecting, an effect called *compound migration*, and therefore affect the performance of the pair.

Table 14-11 shows a common color code for paired cables where they are simply a bundle of pairs. The color coding is only to identify the pair and the coloring of the insulation has no effect on performance. If this cable were individually jacketed pairs, it would be likely that the two wires in the pair would be identical colors such as all black-and-red, and the jackets would use different colors to identify them as shown in Table 14-12.

14.9.2 Analog Multipair Snake Cable

Originally designed for the broadcast industry, hard-wire multipair audio snake cables feature individually shielded pairs, for optimum noise rejection, and sometimes with individual jackets on each pair for improved physical protection. These cables are ideal, carrying multiple line-level or microphone-level signals. They will also interconnect audio components such as multichannel mixers and consoles for recording studios, radio and television stations, postproduction facilities, and sound system installations. Snakes offer the following features:

- A variety insulation materials, for low capacitance, ruggedness, or fire ratings.
- Spiral/serve, braid, French Braid™, or foil shields.
- Jacket and insulation material to meet ruggedness or NEC flame requirements.
- High temperature resistance in some compounds.
- Cold temperature pliability in some compounds.
- Low-profile appearance, based mostly on the gage of the wires, but also on the insulation.

Table 14-11. Color Codes for Paired Cables (Belden Standard)

Pair No.	Color Combination	Pair No.	Color Combination	Pair No.	Color Combination	Pair No.	Color Combination
1	Black/Red	11	Red/Yellow	21	White/Brown	31	Purple/White
2	Black/White	12	Red/Brown	22	White/Orange	32	Purple/Dark Green
3	Black Green	13	Red/Orange	23	Blue/Yellow	33	Purple/Light Blue
4	Black/Blue	14	Green/White	24	Blue/Brown	34	Purple/Yellow
5	Black/Yellow	15	Green/Blue	25	Blue/Orange	35	Purple/Brown
6	Black/Brown	16	Green/Yellow	26	Brown/Yellow	36	Purple/Black
7	Black/Orange	17	Green/Brown	27	Brown/Orange	37	Gray/White
8	Red/White	18	Green/Orange	28	Orange/Yellow		
9	Red/Green	19	White/Blue	29	Purple/Orange		
10	Red/Blue	20	White/Yellow	30	Purple/Red		

Courtesy Belden

- Some feature overall shields to reduce crosstalk and facilitate star grounding.
- Allows easier and cheaper installs than using multiple single channel cables.

Snakes come with various terminations and can be specified to meet the consumer's needs. Common terminations are male or female XLR (microphone) connectors and ¼ inch male stereo connectors on one end, and either a junction box with male or female XLR connectors and ¼ inch stereo connectors or pigtails with female XLR connectors and ¼ inch connectors on the other end.

For stage applications, multipair individually shielded snake cables feature lightweight and small diameter construction, making them ideal for use as portable audio snakes. Individually shielded and jacketed pairs are easier to install with less wiring errors. In areas that subscribe to the NEC guidelines, the need for conduit in studios is eliminated when CM-rated snake cable is used through walls between rooms. Vertically between floors, snakes rated CMR (riser) do not need conduit. In plenum areas (raised floors, drop ceilings) CMP, plenum rated snake cables can be used without conduit. Color codes for snakes are given in Table 14-12.

14.9.3 High Frequency Pairs

Twisted pairs were original conceived to carry low-frequency signals, such as telephone audio. Beginning in the 1970s research and development was producing cables such as twinax that had reasonable performance to the megahertz. IBM Type 1 was the breakthrough product that proved that twisted-pairs could indeed carry data. This led directly to the Category premise/data cable of today.

There are now myriad forms of high-frequency, high-data rate cable including DVI, USB, HDMI, IEEE 1394 FireWire, and others. All of these are commonly used to transport audio and video signals, Table 14-13.

14.9.3.1 DVI

DVI (Digital Visual Interface) is used extensively in the computer-monitor interface market for flat panel LCD monitors.

The DVI connection between local monitors and computers includes a serial digital interface and a parallel interface format, somewhat like combining the broadcast serial digital and parallel digital interfaces.

Transmission of the TMDS (transition minimized differential signaling) format combines four differential, high-speed serial connections transmitted in a parallel bundle. DVI specifications that are extended to the dual mode operation allow for greater data rates for higher display resolutions. This requires seven parallel, differential, high-speed pairs. Quality cabling and connections become extremely important. The nominal DVI cable length limit is 4.5 m (15 ft). Electrical performance requirements are signal rise time of 0.330 ns, and a cable impedance of 100 Ω. FEXT is less than 5%, and signal rise time degradation is a maximum of 160 ps (picoseconds). Cable for DVI is application specific since the actual bit rate per channel is 1.65 Gbps.

Table 14-12. Color Codes for Snake Cables

Pair No.	Color Combination	Pair No.	Color Combination	Pair No.	Color Combination	Pair No.	Color Combination	Pair No.	Color Combination
1	Brown	13	Lt. Gray/Brown stripe	25	Lt. Blue/Brown stripe	37	Lime/Brown stripe	49	Aqua/Brown stripe
2	Red	14	Lt. Gray/Red stripe	26	Lt. Blue/Red stripe	38	Lime/Red stripe	50	Aqua/Red stripe
3	Orange	15	Lt. Gray/Orange stripe	27	Lt. Blue/Orange stripe	39	Lime/Orange stripe	51	Aqua/Orange stripe
4	Yellow	16	Lt. Gray/Yellow stripe	28	Lt. Blue/Yellow stripe	40	Lime/Yellow stripe	52	Aqua/Yellow stripe
5	Green	17	Lt. Gray/Green stripe	29	Lt. Blue/Green stripe	41	Lime/Green stripe	53	Aqua/Green stripe
6	Blue	18	Lt. Gray/Blue stripe	30	Lt. Blue/Blue stripe	42	Lime/Blue stripe	54	Aqua/Blue stripe
7	Violet	19	Lt. Gray/Violet stripe	31	Lt. Blue/Violet stripe	43	Lime/Violet stripe	55	Aqua/Violet stripe
8	Gray	20	Lt. Gray/Gray stripe	32	Lt. Blue/Gray stripe	44	Lime/Gray stripe	56	Aqua/Gray stripe
9	White	21	Lt. Gray/White stripe	33	Lt. Blue/White stripe	45	Lime/White stripe	57	Aqua/White stripe
10	Black	22	Lt. Gray/Black stripe	34	Lt. Blue/Black stripe	46	Lime/Black stripe	58	Aqua/Black stripe
11	Tan	23	Lt. Gray/Tan stripe	35	Lt. Blue/Tan stripe	47	Lime/Tan stripe	59	Aqua/Tan stripe
12	Pink	24	Lt. Gray/Pink stripe	36	Lt. Blue/Pink stripe	48	Lime/Pink stripe	60	Aqua/Pink stripe

Courtesy Belden.

Table 14-13. Comparing Twisted-Pair High-Frequency Formats

Standard	Format	Intended Use	Connector Style	Cable Type	Transmission Distance[1]	Sample Rate	Data Rate (Mbps)	Guiding Document
D1 component	parallel	broadcast	multipin D	multipairs	4.5 m/15 ft	27 MHz	270	ITU-R BT.601-5
DV	serial	professional/ consumer		(see IEEE 1394)	4.5 m/15 ft	20.25 MHz	25	IEC 61834
IEEE 1394 (FireWire)	serial	professional/ consumer	1394	6 conductors, 2-STPs/2 pwr	4.5 m/15 ft	n/a	100, 200, 400	IEEE 1394
USB 1.1	serial	consumer	USB A & B	4 conductors, 1 UTP/ 2 pwr	5 m/16.5 ft	n/a	12	USB 1.1 Promoter Group
USB 2.0	serial	professional/ consumer	USB A & B	4 conductors, 1-UTP/ 2 pwr	5 m/16.5 ft	n/a	480	USB 2.0 Promoter Group
DVI	serial/ parallel	consumer	DVI (multipin D)	Four STPs	10 m/33 ft	To 165 MHz	1650	DDWG; DVI 1.0
HDMI	parallel	consumer	HDMI (19 pin)	Four STPs + 7 conductors	Unspecified	To 340 MHz	To 10.2 Gbps	HDMI LLC
DisplayPort	parallel	consumer	20 pin	Four STPs + 8 conductors	15 m	To 340 MHz	To 10.8 Gbps	VESA

[1] Transmission distances may vary widely depending on cabling and the specific equipment involved.
STP = shielded twisted pair, UTP = unshielded twisted pair, n/a = not applicable

Picture information or even the entire picture can be lost if any vital data is missing with digital video interfaces. DVI cable and its termination are very important and the physical parameters of the twisted pairs must be highly controlled as the specifications for the cable and the receiver are given in fractions of bit transmission.

Requirements depend on the clock rate or signal resolution being used. Transferring the maximum rate of 1600 × 1200 at 60 Hz for a single link system means that one bit time or 10 bits per pixel is 0.1(1 ⁄ 165 MHz) or 0.606 ns.

The DVI receiver specification allows 0.40 × bit time, or 0.242 ns intrapair skew within any twisted pair. The pattern at the receiver must be very symmetrical. The interpair skew, which governs how bits will line up in time at the receiving decoder, may only be 0.6 × pixel time, or 3.64 ns. These parameters control the transmission distances for DVI.

Also, the cable should be evaluated on its insertion loss for a given length. DVI transmitter output is specified into a cable impedance of 100 Ω with a signal swing of ±780 mV with a minimum signal swing of ±200 mV. When determining DVI cable, assume minimum performance by the transmitter—i.e., 200 mV—and best sensitivity by the receiver which must operate on signals ±75 mV. Under these conditions the cable attenuation can be no greater than 8.5 dB at 1.65 GHz (10 bits/pixel × 165 MHz clock) which is relatively difficult to maintain on twisted-pair cable.

DVI connections combine the digital delivery, described above, with legacy analog component delivery. This allows DVI to be the transition delivery scheme between analog and digital applications.

14.9.3.2 HDMI

HDMI (high definition multimedia interface) is similar to DVI except that it is digital-only delivery. Where DVI has found its way into the commercial space as well as consumer applications, HDMI is almost entirely consumer-based. It is configured into a 19 pin connector which contains four shielded twisted pairs (three pairs data, one pair clock) and seven wire for HDCP (copy protection), devices handshaking, and power. The standard versions of HDMI are nonlocking connector, attesting to its consumer-only focus.

14.9.3.3 IEEE -1394 or FireWire Serial Digital

FireWire, or IEEE -1394, is used to upload DV, or digital video, format signals to computers etc. DV, sometimes called DV25, is a serial digital format of 25 Mbps. IEEE 1394 supports up to 400 Mbps. The specification defines three signaling rates, S100 (98.304 Mbps), S200 (196.608 Mbps), and S400 (393.216 Mbps).

IEEE 1394 can interconnect up to sixty three devices in a peer-to-peer configuration so audio and video can be transferred from device to device without a computer, D/A, or A/D conversion. IEEE 1394 is hot plugable from the circuit while the equipment is turned on.

The IEEE 1394 system uses two shielded twisted pairs and two single wires, all enclosed in a shield and jacket, Fig. 14-2. Each pair is shielded with 100% coverage foil and a minimum 60% coverage braid. The outer shield is 100% coverage foil and a minimum 90% coverage braid. Each pair is shielded with aluminum foil and is equal to or greater than 60% braid. The twisted pairs handle the differential data and strobe (assists in clock regeneration) while the two separate wires provide the power and ground for remote devices. Signal level is 265 mV differential into 110 Ω.

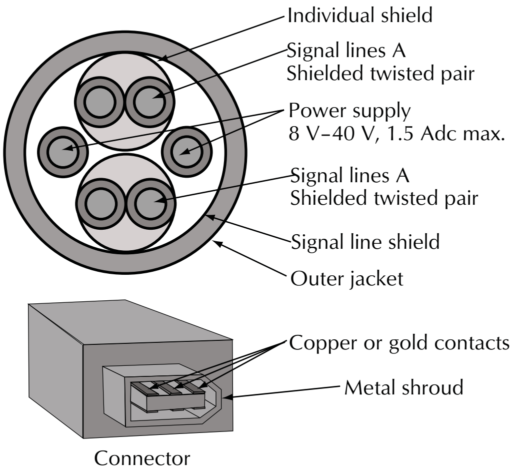


Figure 14-2. IEEE 1394 cable and connector.

The IEEE 1394 specification cable length is a maximum of 4.5 m (15 ft). Some applications may run longer lengths when the data rate is lowered to the 100 Mbps level. The typical cable has #28 gage copper twisted pairs and #22 gage wires for power and ground. The IEEE 1394 specification provides the following electrical performance requirements:

- Pair-to-pair data skew is 0.40 ns.
- Crosstalk must be maintained below −26 dB from 1 MHz to 500 MHz.
- Velocity of propagation must not exceed 5.05 ns/m.

Table 14-14 gives details of the physical interface system for IEEE 1394.

14.9.3.4 USB

The USB, universal serial bus, simplifies connection of computer peripherals. USB 1.1 is limited to a communications rate of 12 Mbps, while USB 2.0 supports up to 480 Mbps communication. The USB cable consists of one twisted pair for data and two untwisted wires for powering downstream appliances. A full-speed cable includes a #28 gage twisted pair, and an untwisted pair of #28 to #20 gage power conductors, all enclosed in an aluminized polyester shield with a drain wire.

Table 14-14. Critical IEEE 1394 Timing Parameters

Parameter	100 Mbps	200 Mbps	400 Mbps
Max *Tr*/*Tf*	3.20 ns	2.20 ns	1.20 ns
Bit Cell Time	10.17 ns	5.09 ns	2.54 ns
Transmit Skew	0.40 ns	0.25 ns	0.20 ns
Transmit Jitter	0.80 ns	0.50 ns	0.25 ns
Receive End Skew	0.80 ns	0.65 ns	0.60 ns
Receive End Jitter	1.08 ns	0.75 ns	0.48 ns

Nominal impedance for the data pair is 90 Ω. The maximum cable length is determined by the signal propagation delay which must be less than 26 ns from end to end. Table 14-15 lists some common plastics and the theoretical distance each could go based on 26 ns. With an additional allowance of 4 ns, which is split between the sending device connection and the receiver connection/response function, the entire one-way delay is a maximum of 30 ns. The cable velocity of propagation must be less than 5.2 ns/m and the length and twist of the data pair must be matched so time skew is no more than 0.10 ns between bit polarities. The nominal differential signal level is 800 mV.

Table 14-15. Dielectric Constant, Delay, and Transmission Distance of Various Plastics

Material	Dielectric Constant	Delay ns/ft	Maximum USB Distance
Foam, Air-Filled Plastic	1.35	1.16	22.4 ft
Solid Teflon™	2.1	1.45	18 ft
Solid Polyethylene	2.3	1.52	17 ft
Solid Polyvinyl Chloride	3.5–6.5	1.87–2.55	10–14 ft

14.9.3.5 DisplayPort

DisplayPort is an emerging protocol for digital video. Its original intention was the transfer of images from a PC or similar device to a display. It has some significant advantages over DVI and HDMI. DisplayPort is by design backward-compatible to single link DVI and HDMI. Those are both severely distance-limited by the delay skew of the three data pairs when compared to the clock pair. With DisplayPort the clock is embedded with the video, much as the clock is embedded with the audio bit stream in AES digital audio, so the distance limitations on DisplayPort are less likely to involve clock timing problems.

However, display port is also a nonlocking connector, of 20 pins, and is intended for maximum distance 15 m (50 ft). These cables are, like HDMI and DVI, only available in assemblies. Raw cable and connectorization in the field do not currently look like an option for the professional installer. All these factors make it less likely to be embraced by the professional broadcast video arena.

14.9.3.6 Premise/Data Category Cables

While premise/data category cables were never intended to be audio or video cables, their high performance and low cost, and their ubiquitous availability, have seen them pressed into service carrying all sorts of nondata signals.

It should also be noted that high-speed Ethernet networks are routinely used to transport these audio and video signals in data networks. The emergence of 10GBase-T, 10 gigabit networks, will allow the transport of even multiple uncompressed 1080p/60 video images. The digital nature of most entertainment content, with the ubiquitous video server technology in use today, makes high-bandwidth, high-data-rate networks in audio, video, broadcast, and other entertainment facilities, an obvious conclusion.

14.9.3.6.1 Cabling Definitions

- *Telcom Closet* (TC). Location where the horizontal cabling and backbone cabling are made.
- *Main Cross-Connect* (MXC). Often called the equipment room and is where the main electronics are located.
- *Intermediate Cross-Connect* (IXC). A room between the TC and the MXC are terminated. Rarely used in LANs.
- *Horizontal Cabling*. The connection from the telcom closet to the work area.
- *Backbone Cabling*. The cabling that connect all of the hubs together.
- *Hub*. The connecting electronic box that all of the horizontal cables connect to which are then connected to the backbone cable.
- *Ethernet*. A 10, 100, or 1000 Mb/s LAN. The 10 Mbps version is called 10Base-T. The 100 Mbps version is called Fast Ethernet and 1000 Mbps version is called Gigabit Ethernet.

14.9.3.6.2 Structured Cabling

Structured cabling, also called communications cabling, data/voice, low voltage, or limited energy is the standardized infrastructure for telephone and local area network (LAN) connections in most commercial installations. The architecture for the cable is standardized by Electronic Industries Association and Telecommunications Industry Association (EIA/TIA), an industry trade association. EIA/TIA 568, referred to as 568, is the main document covering structured cabling. IEEE 802.3 also has standards for structured cabling.

The current standard, as of this writing, is EIA/TIA 568-B.2-10 that covers all active standards up to 10GbaseT, 10 gigabit cabling.

14.9.3.6.3 Types of Structured Cables

Following are the types of cabling, Category 1 though Category 7, often referred to as Cat 1 through Cat 7. The standard TIA/EIA 568A no longer recognizes Cat 1, 2, or 4. As of July 2000, the FCC mandated the use of cable no less than Cat 3 for home wiring. The naming convention specified by ISO/IEC 11801 is shown in Fig. 14-3.

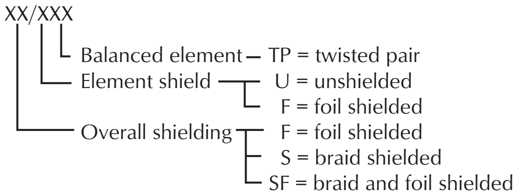


Figure 14-3. ISO/IEC 11801 cable naming convention.

Table 14-16 gives the equivalent TIA and ISO classifications for structured cabling.

Table 14-16. TIA and ISO Equivalent Classifications

Frequency bandwidth	TIA		ISO	
	Components	Cabling	Components	Cabling
1–100 MHz	Cat 5e	Cat 5e	Cat 5e	Class D
1–250 MHz	Cat 6	Cat 6	Cat 6	Class E
1–500 MHz	Cat 6a	Cat 6a	Cat 6a	Class E_A
1–600 MHz	n/s	n/s	Cat 7	Class F
1–1000 MHz	n/s	n/s	Cat 7A	Class F_A

Category 1. Meets the minimum requirements for analog voice or plain old telephone service (POTS). This category is not part of the EIA/TIA 568 standard.

Category 2. Defined as the IBM Type 3 cabling system. IBM Type 3 components were designed as a higher grade 100 Ω UTP system capable of operating 1 Mb/s Token Ring, 5250, and 3270 applications over shortened distances. This category is not part of the EIA/TIA 568 standard.

Category 3. Characterized to 16 MHz and supports applications up to 10 Mbps. Cat 3 conductors are 24 AWG. Applications range from voice to 10Base-T.

Category 4. Characterized to 20 MHz and supports applications up to 16 Mb/s. Cat 4 conductors are 24 AWG. Applications range from voice to 16 Mbps Token Ring. This category is no longer part of the EIA/TIA 568 standard.

Category 5. Characterized to 100 MHz and supports applications up to 100 Mbps. Cat 5 conductors are 24 AWG. Applications range from voice to 100Base-T. This category is no longer part of the EIA/TIA 568 standard.

Category 5e. Characterized to 100 MHz and supports applications up to 1000 Mbps/1 Gbps. Cat 5e conductors are 24 AWG. Applications range from voice to 1000Base-T. Cat 5e is specified under the TIA standard ANSI/TIA/EIA-568-B.2. Class D is specified under ISO standard ISO/IEC 11801, 2nd Ed.

Category 6. Characterized to 250 MHz, in some versions bandwidth is extended to 600 MHz, and supports 1000 Mbps/1 Gbps and future applications and is backward compatible with Cat 5 cabling systems. Cat 6 conductors are 23 AWG. This gives improvements in power handling, insertion loss, and high-frequency attenuation. Fig. 14-4 shows the improvements of Cat 6 over Cat 5e. Cat 6 is specified under the TIA standard ANSI/TIA/EIA-568-B.2-1. Class E is specified under ISO standard ISO/IEC 11801, 2nd Ed. Cat 6 is available most commonly in the United States as UTP.

Category 6 F/UTP. Cat 6 F/UTP (foiled unshielded twisted pair) or ScTP (screened twisted pair) consists of four twisted pairs enclosed in a foil shield with a conductive material on one side. A drain wire runs adjacent to the conductive side of the shield, Fig. 14-5. When appropriately connected, the shield reduces ANEXT, RFI, and EMI. Cat 6 FTP can only be designed to 250 MHz per TIA/EIA 568B.2-1.

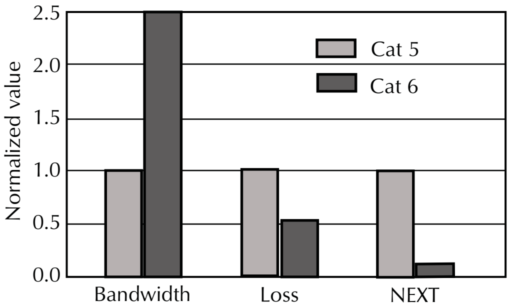
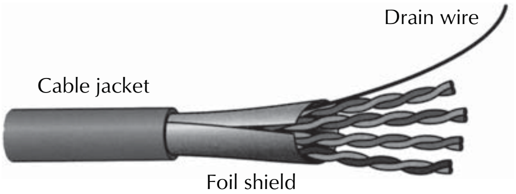


Figure 14-4. Normalized comparison of Cat 5e and Cat 6.



Source: BICSI

Figure 14-5. Cat 6 F/UTP.

Category 6a. Cat 6a (Augmented Category 6) is characterized to 500 MHz, and in special versions to 625 MHz, has lower insertion loss, and has more immunity to noise. Cat 6a is often larger than the other cables. 10GBase-T transmission uses digital signal processing (DSP) to cancel out some of the internal noise created by NEXT and FEXT between pairs. Cat 6a is specified under the TIA standard ANSI/TIA/EIA 568-B.2-10. Class E_A is specified under ISO standard ISO/IEC 11801, 2nd Ed. Amendment 1. Cat 6a is available as UTP or FTP.

Category 7 S/STP. Cat 7 S/STP (foil shielded twisted-pair) cable is sometimes called PiMF (pairs in metal foil). Shielded-twisted pair 10GBase-T cable dramatically reduces alien crosstalk. Shielding reduces electromagnetic interference (EMI) and radio-frequency interference (RFI). This is particularly important as the airways are getting more congested. The shield reduces signal leakage and makes it harder to tap by an outside source. Shield termination at 14.16 Class F will be specified under ISO standard ISO/IEC 11801, 2nd Ed. Class FA will be specified under ISO standard ISO/IEC 11801, 2nd Ed. Amendment 1.

14.9.3.6.4 Comparisons

Table 14-17 compares network data rates for Cat 3 through Cat 6a and Table 14-18 compares various characteristics of Cat 5e, 6, and 6a. Fig. 14-6 compares the media distance-bandwidth product of Cat 5e and Cat 6a with 802.11 (a, b, g, n) wireless media, often called Wi-Fi.

Table 14-17. Network Data Rates, Supporting Cable Types, and Distance

Minimum Performance	Token Ring	Ethernet	Maximum Distance
Cat 3	4 Mb/s	10 Mbps	100 m/328 ft
Cat 4	16 Mb/s	–	100 m/328 ft
Cat 5	–	100 Mbps	100 m/328 ft
Cat 5e		1000 Mbps	100 m/328 ft
Cat 6	–	10 Gbps	55 m/180 ft
Cat 6a		10 Gbps	100 m/328 ft

Table 14-18. Characteristics of Cat 5e, Cat 6, and Cat 6a

Cabling Type	Cat 5e	Cat 6	Cat 6a
Relative Price (%)	100	135–150	165–180
Available Bandwidth	100 MHz	250 MHz	500 MHz
Data rate Capability	1.2 Gbps	2.4 Gbps	10 Gbps
Noise Reduction	1.0	0.5	0.3
Broadband Video Channels 6 MHz/channel	17	42	83
Broadband Video Channels rebroadcast existing channels	6	28	60+
No. of Cables in Pathway 24 inches × 4 inches	1400	1000	700

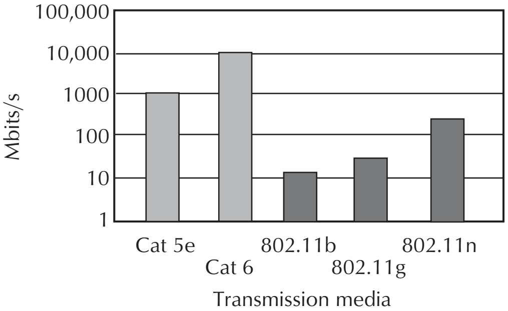


Figure 14-6. Comparison of media distance to bandwidth

New cable designs can affect size and pathway load so consult the manufacturer. Note that cable density is continually changing with newer, smaller cable designs. Numbers in Table 14-18 should be considered worst case. Designers and installers of larger systems should get specific dimensional information from the manufacturer.

Fig. 14-7 shows various problems that can be found in UTP cabling. Fig. 14-8 gives the maximum distances for UTP cabling as specified by ANSI/TIA.

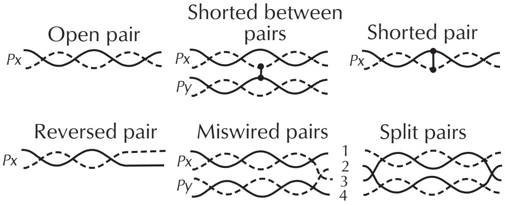


Figure 14-7. Paired wiring faults. Courtesy Belden.

Four (4) pair 100 Ω ±15% UTP Cat 5e cabling is the recommended minimum requirement for residential and light commercial installations because it provides excellent flexibility. Pair counts are four pair for desktop and twenty five pair for backbone cabling. The maximum length of cable is 295 ft (90 m) with another 33 ft (10 m) for patch cords.

Unshielded twisted pairs (UTP) and shielded twisted pairs (STP) are used for structured cabling. Unshielded twisted pairs (UTP) are the most common today. These cables look like the POTS cable, however, their construction makes them usable in noisy areas and at high frequencies because of the short, even twisting of the two wires in each pair. The twist must be even and tight so complete noise cancellation occurs along the entire length of the cable. To best keep the twist tight and even, better cable has the two wires bonded together so they will not separate when bent or flexed. Patch cable is flexible so twist and impedance are not as well controlled. The color codes for the pairs are given in Table 14-19.

Cable diameter varies for the different types of cable. TIA recommends that two Cat 6 cables but only one Cat 6a cable can be put in a ¾ inch (21 mm) conduit at 40% fill. The diameter and the stiffness of the cables determine their bend radius and therefore the bend radius of conduits and trays, Table 14-20

Fig. 14-9 shows the construction of UTP and screened UTP cable.

14.9.3.6.5 Critical Parameters

Critical parameters for UTP cable are: NEXT, PS-NEXT, FEXT, ELFEXT, PS-ELFEXT, RL, ANEXT.

NEXT. *NEXT*, or near-end crosstalk, is the unwanted signal coupling from the near end of one sending pair to a receiving pair.

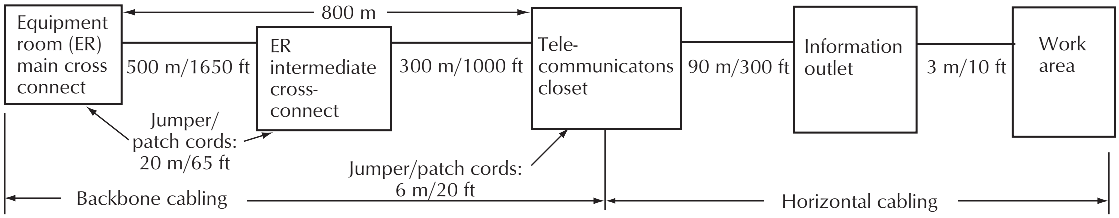

800 m
Equipment room (ER) main cross connect
500 m/1650 ft
ER intermediate cross-connect
300 m/1000 ft
Tele-communicatons closet
90 m/300 ft
Information outlet
3 m/10 ft
Work area
Jumper/ patch cords: 20 m/65 ft
Jumper/patch cords: 6 m/20 ft
Backbone cabling
Horizontal cabling

Figure 14-8. Maximum distances between areas for UTP cable.

Table 14-19. Color Code for UTP Cable

Pair No.	1st Conductor Base/Band	2nd Conductor
1	White/Blue	Blue
2	White/Orange	Orange
3	White/Green	Green
4	White/Brown	Brown

Table 14-20. Diameter and Bend Radius for 10GbE Cabling

Cable	Diameter	Bend Radius
Category 6	0.22 inch (5.72 mm)	1.00 inch (4 × OD)
Category 6a	0.35 inch (9 mm)	1.42 inch (4 × OD)
Category 6 FTP	0.28 inch (7.24 mm)	2.28 inch (8 × OD)
Category 7 STP	0.33 inch (8.38 mm)	2.64 inch (8 × OD)

Cable diameters are nominal values.

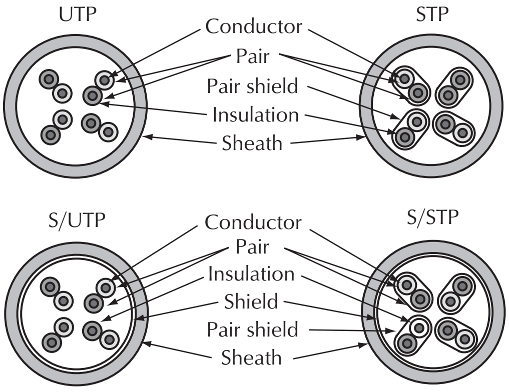

UTP
STP
Conductor
Pair
Pair shield
Insulation
Sheath
S/UTP
S/STP
Conductor
Pair
Insulation
Shield
Pair shield
Sheath

Figure 14-9. UTP and S/UTP cable design.

PS-NEXT. *PS-NEXT*, or power-sum near-end crosstalk, is the crosstalk between all of the sending pairs to a receiving pair. With four-pair cable, this is more important than NEXT.

FEXT. *FEXT*, or far-end crosstalk, is the measure of the unwanted signal from the transmitter at the near end coupling into a pair at the far end.

EL-FEXT. *EL-FEXT*, or equal level far-end crosstalk, is the measure of the unwanted signal from the transmitter end to a neighboring pair at the far end relative to the received signal at the far end. The equation is

$$EL-FEXT = FEXT-Attenuation \tag{14-3}$$

Power sum equal-level far-end crosstalk is the computation of the unwanted signal coupling from multiple transmitters at the near end into a pair measured at the far end relative to the received signal level measured on the same pair.

Return Loss (RL). *RL* is a measure of the reflected energy from a transmitted signal and is expressed in –dB, the higher the value, the better. The reflections are caused by impedance mismatch caused by connectors, improper installation such as stretching the cable or too sharp a bend radius, improper manufacturing, or improper load.

Broadcasters are very familiar with return loss, calling it by a different name, SWR (standing wave ratio) or VSWR (voltage standing wave ratio). In fact, return loss measurements can easily be converted into VSWR values, or vice versa. Return loss can be found with the equation

$$RL = 20\log\frac{Difference}{Sum} \tag{14-4}$$

where,

Difference is the difference (absolute value) between the desired impedance and the actual measured impedance,

Sum is the desired impedance and the actual measured impedance added together.

- The desired impedance for all UTP data cables (Cat 5, 5e, 6, 6a) is 100 Ω.
- The desired impedance for all passive video, HD, HD-SDI or 1080p/60 components is 75 Ω.
- The desired impedance for all digital audio twisted pairs is 110 Ω.

- The desired impedance for all digital audio on coaxial cable is 75 Ω.

With 1000Base-T systems, the pairs simultaneously transmit and receive. As the transmitter sends data, it is also listening for data being sent from the opposite end of the same pair. Any reflected signal from the sending end that reflects back to the sending end mixes with the sending signal from the far end, reducing intelligibility. With 10Base-T or 100Base-T data networks, one pair transmits while another receives, so reflections (RL, return loss) are not a major consideration and were not required to be measured. Now, with pairs simultaneously transmitting and receiving, called duplex mode, RL is a critical measurement for data applications.

Delay Skew. Since every pair (and every cable) takes a specific amount of time to deliver a signal from one end to the other, there is a delay. Where four pairs must deliver data to be recombined, the delay in each pair should, ideally, be the same. However, to reduce crosstalk, the individual pairs in any Category cable have different twist rates. This reduces the pair-to-pair crosstalk but affects the delivery time of the separate parts. This is called *delay skew*.

While delay skew affects the recombining of data, in 1000Base-T systems, for instance, the same delay skew creates a problem when these UTP data cables are used to transmit component video or similar signals, since the three colors do not arrive at the receiving end at the same time, creating a thin bright line on the edge of dark images. Some active baluns have skew correction built in, see Section 14.12.5.

ANEXT. *ANEXT*, or alien crosstalk, is coupling of signals between cables. This type of crosstalk cannot be cancelled by DSP at the switch level. Alien crosstalk can be reduced by overall shielding of the pairs, or by inserting a nonconducting element inside that cable to push away the cables around it.

14.9.3.6.6 Terminating Connectors

All structured cabling use the same connector, an RJ-45. In LANs (local area networks) there are two possible pin-outs, 568A and 568B. The difference is pair 2 and pair 3 are reversed. Both work equally well as long as they are not intermixed. The termination is shown in Fig. 14-10.

In the past decade, the B wiring scheme has become the most common. However, if you are adding to or extending and existing network, you must determine which wiring scheme was used and continue with that scheme. A mixed network is among the most common causes of network failure.

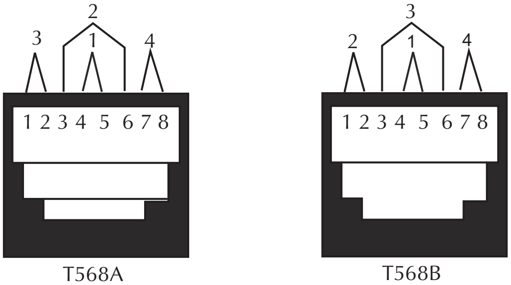


Figure 14-10. Termination layout for EIA/TIA 568-B.2 cable.

It is very important that the pairs be kept twisted as close to the connector as possible. For 100Base-T (100 MHz, 100 Mbps) applications, a maximum of ½ inch should be untwisted to reduce crosstalk and noise pickup. In fact, with Cat 6 (250 MHz) or Cat 6a (500 MHz) it is safe to say that any untwisting of the pairs will affect performance. Therefore there are many connectors, patch panels, and punch-down blocks that minimize the untwisting of the pairs.

14.9.3.6.7 Baluns

Baluns (*Bal*anced-*Un*balanced) networks are a method of connecting devices of different impedance and different formats. Baluns have been commonly used to convert unbalanced coax, to balanced twinlead for television antennas, or to match coaxial data formats (coaxial Ethernet) to balanced line systems (10Base-T, 100Base-T etc.). Other balun designs can allow unbalanced sources, such as video or consumer audio, for instance, to be carried on balanced lines, such as UTP Cat 5e, 6, etc.

Since there are four pairs in a common data cable, this can carry four channels. Since category cables are rarely tested below 1 MHz, the audio performance was originally suspect. Crosstalk at audio frequencies in UTP has been measured and is consistently better than –90 dB even on marginal Cat 5. On Cat 6, the crosstalk at audio frequencies is below the noise floor of most network analyzers.

Baluns are commonly available to handle such signals as analog and digital audio, composite video, S-video, RGB or other component video (VGA, Y/R-Y/B-Y,Y/Cr/Cb), broadband RF/CATV, and even DVI and HDMI. The limitations to such applications are the bandwidth specified on the cable and the performance of the cable (attenuation, return loss, crosstalk, etc.) at those higher frequencies.

Passive baluns can also change the source impedance in audio devices. This dramatically extends the effective distance of such signals from only a few feet to many hundreds of feet. Consult the balun manufacturer for the actual output impedance of their designs.

Some baluns can include active amplification, equalizations, or skew (delivery timing) compensation. While more expensive, these active baluns can dramatically increase the effective distance of even marginal cable.

14.9.3.6.8 Adaptors

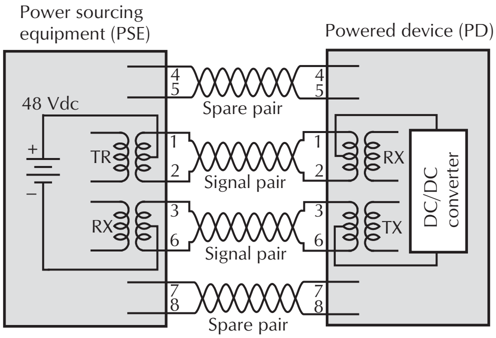
Users and installers should be aware there are adaptors, often that fit in wall plates, where keystone data jacks are intended to be snapped in place. These adaptors often connect consumer audio and video (RCA connectors) to 110 blocks or other twisted pair connection points. However, there is no unbalanced-to-balanced device in these, so the noise rejection inherent in twisted pairs when run as a balanced line is not provided. These adaptors simply unbalance the twisted pair and offer dramatically short effective distances. Further, baluns can change the source impedance and extend distance. Adaptors with no transformers or similar components cannot extend distance and often reduce the effective distance. These devices should be avoided unless they contain an actual balun.

14.9.3.6.9 Power Over Ethernet (PoE)

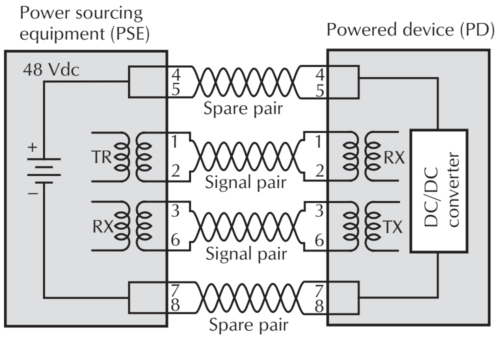
PoE supplies power to various Ethernet services as VoIP (Voice over Internet Protocol) telephones, wireless LAN access points, Blu-tooth access points, and Web cameras. Many audio and video applications will soon use this elegant powering system. IEEE 802.3af-2003 is the IEEE standard for PoE. IEEE 802.3af specifies a maximum power level of 15.4 W at the power sourcing equipment (PSE) and a maximum of 12.95 W of power over two pairs to a powered device (PD) at the end of a 100 m (330 ft) cable.

The PSE can provide power by one of two configurations:

1. Alternative A, sometimes called phantom powering, supplies the power over pairs 2 and 3.
2. Alternative B supplies power over pairs 1 and 4, as shown in Fig. 14-11.



A. Power supplied via data pairs



B. Power supplied by spare pairs

Figure 14-11. Two methods of supplying power via the Ethernet. Courtesy of Panduit Int'l Corp.

The voltage supplied is nominally 48 Vdc with a minimum of 44 Vdc, a maximum of 57 Vdc, and the maximum current per pair is 350 mAdc, or 175 mAdc per conductor. For a single solid 24 AWG wire, common to many category cable designs, of 100 m length (328 ft) this would be a resistance of 8.4 Ω. Each conductor would dissipate 0.257 W or 1.028 W per cable (0.257 W × 4 conductors). This causes a temperature rise in the cable and conduit which must be taken into consideration when installing PoE.

14.9.3.6.10 Power Over Ethernet Plus (PoE Plus)

PoE Plus is defined in IEEE 802.3at and is capable of delivering up to 30 W. Work is being done to approach 60 W or even greater. This requires the voltage supply to be 50 to 57 Vdc. Assuming a requirement of 42 W of power at the endpoint at 50 Vdc, the total current would be 0.84 A, or 0.21 A per pair, or 0.105 A (105 mA) per conductor, or a voltage drop of only 0.88 V in one 24 AWG wire.

14.10 Coaxial Cable

Coaxial cable is a design in which one conductor is accurately centered inside another with both conductors carrying the desired signal currents (source to load and return), as shown in Fig. 14-12. Coaxial cable is so called because if you draw a line through the center of a cross-sectional view, you will dissect all parts of the cable. All parts are on the same axis, or coaxial.

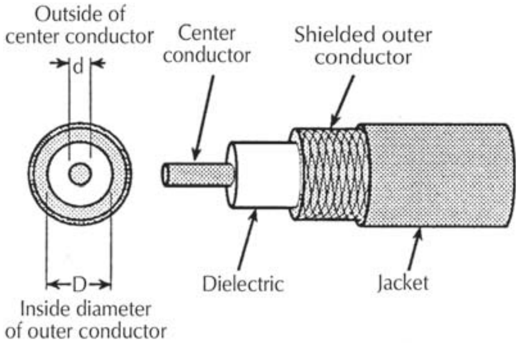

Outside of center conductor
d
Center conductor
Shielded outer conductor
D
Inside diameter of outer conductor
Dielectric
Jacket

Figure 14-12. Construction of a coaxial cable.

14.10.1 History of Coaxial Cable

It has been argued that the first submarine telegraph cable (1858) was coaxial, Fig. 14-13. While this did have multiple layers, the outer layer was not part of the signal-carrying portion. It was a protective layer.

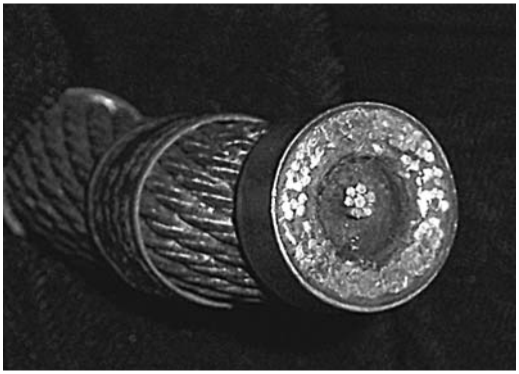


Figure 14-13. First submarine cable.

Modern coaxial cable was invented on May 23, 1929 by Lloyd Espenscheid and Herman Affel of Bell Laboratories. Often called *coax*, it is often used for the transmission of high-frequency signals. At high frequencies, above 100 kHz, coax has a dramatically better performance than twisted pairs. However, coax lacks the ability to reject noise that twisted pairs can do when configured as balanced lines. Coaxial cable was first installed in 1931 to carry multiple telephone signals between cities.

14.10.2 Coaxial Cable Construction

The insulation between the center conductor and the shield of a coaxial cable affects the impedance and the durability of the cable. The best insulation to use between the center conductor and the shield would be a vacuum. The second best insulation would be dry air, the third, nitrogen. The latter two are familiar insulators in hard-line transmission line commonly used to feed high-power antenna in broadcasting.

A vacuum is not used, even though it has the lowest dielectric constant of “1,” because there would be no conduction of heat from the center conductor to the outer conductor and such a transmission line would soon fail. Air and nitrogen are commonly used under pressure in such transmission lines. Air is occasionally used in smaller, flexible cables.

Polyethylene (PE) was common as the core material in coaxial cables during WW II. Shortly after the war, polyethylene was declassified and most early cable designs featured this plastic. Today most high-frequency coaxial cables have a chemically formed foam insulation or a nitrogen gas injected foam. The ideal foam is high-density hard cell foam, which approaches the density of solid plastic but has a high percentage of nitrogen gas. Current state-of-the-art polyethylene foam velocity is 86% (dielectric constant: 1.35) although most digital video cables are 82–84% velocity of propagation. High-density foam of this velocity resists conductor migration when the cable is bent, keeping impedance variations to a minimum. This high velocity improves the high-frequency response of the cable.

A problem with soft foam is it easily deforms, which changes the distance between the center conductor and the shield, changing the cable impedance. This can be caused by bending the cable too sharply, or running over it, or pulling it too hard, or any other possibility. To reduce this problem, a hard cell foam is used. Some cable that is rated as having a very high velocity of propagation might use very soft foam. A simple test can be performed where the user squeezes the foam dielectric of various cables. It will be immediately apparent that some cables have a density (crush resistance) double that of other designs. Soft foam can lead to conductor migration over time which will change timing, impedance, return loss, and bit errors over distance.

Coaxial cable is used quite extensively with various types of test equipment. When such cable is replaced, the capacitance per foot, which is determined by the dielectric constant of the insulator, must be taken into consideration, particularly for oscilloscope probes.

14.10.2.1 CCTV Cable

CCTV (closed circuit television) cable has a 75 Ω characteristic impedance. CCTV is a baseband signal comprised of low-frequency vertical and horizontal sync pulse information and high-frequency video information. Since the signal is broadband, only cable with a center conductor of solid copper should be used.

If the cable is constantly in motion as in pan and tilt operation, a stranded center conductor should be used as a solid conductor will work-harden and break. There are also robotic coaxes designed to flex millions of times before failure for intense flexing applications.

Shielding for CCTV cable should have a copper or tinned-copper braid of at least 80% coverage, for low-frequency noise rejection. If an aluminum foil shield is used in conjunction with a braid, either tinned copper or aluminum only may be used for the shield. A bare copper braid will result in a galvanic reaction.

14.10.2.1.1 CCTV Distances

For common CCTV 75 Ω cables, their rule-of-thumb transmission distances are shown in Table 14-21. These distances can be extended by the use of in-line booster amplifiers.

Table 14-21. Transmission Distances for CCTV Cable

RG-59	1000 ft
RG-6	1500 ft
RG-11	3000 ft

14.10.2.2 CATV Broadband Cable

For higher-frequency applications, such as carrying radio frequencies or television channels, only the skin of the conductor is working (see Section 14.2.8, Skin Effect). Television frequencies in the United States, for instance, start with Channel 2 (54 MHz) which is definitely in the area of skin effect. So these cables can use center conductors that have a layer of copper over a steel wire, since only the copper layer will be working.

If one uses a copper-clad steel conductor for applications below 50 MHz, the conductor has a dc resistance from four to seven times that of a solid copper conductor. If a copper-clad cable is used on a baseband video signal, for instance, the sync pulses may be attenuated too much. If such a cable is used to carry audio, almost the entire audio signal will be running down the steel wire.

CATV/broadband cable should have a foil shield for good high-frequency noise rejection. CATV cable should also have a braid shield to give the connector something to grab onto, 40% to 60% aluminum braid being the most common. Multiple layer shields are also available such as tri-shielded (foil-braid-foil) and quad shields (foil-braid-foil-braid). Assumptions that quad shields give the best shield effectiveness are erroneous, there being single foil/braid and tri-shield configurations that are measurably superior. Refer to Section 14.8.6 on shield effectiveness.

Modern CATV/broadband cable will use a foamed polyethylene or foamed FEP dielectric, and preferably one with gas injected foam. This will reduce the losses in the cable. The jacket material is determined by the environment that the cable will be working in (see Sections 14.4, 14.5, 14.6).

14.10.3 Coaxial Cable Installation Considerations

14.10.3.1 Indoor Installation

Indoor environments are the most common for coaxial cable installations. A few tips on installing coaxial cable are as follows:

1. First and foremost, follow all NEC requirements when installing coaxial cables.
2. Distribute the pulling tension evenly over the cable and do not exceed the minimum bend radius of ten times the diameter. Exceeding the maximum pulling tension or the minimum bend radius of a cable can cause permanent damage both mechanically and electrically to the cable.
3. When pulling cable through conduit, clean and deburr the conduit completely and use proper lubricants in long runs.

14.10.3.2 Outdoor Installation

Outdoor installations require special installation techniques that will enable the cable to withstand harsh environments. When using cable in an aerial application, lash the cable to a steel messenger, or buy cable with a built-in steel messenger. This will help support the cable and reduce the stress on the cable during wind, snow, and ice storms. When direct burying a cable, lay the cable without tension so it will not be stressed when earth is packed around it. When burying in rocky soil, fill the trench with sand. Lay the cable and then place pressure-treated wood or metal plates over the cable. This will prevent damage to the cable from rocky soil settling; in cold climate areas, bury the cable below the frost line. Buy direct burial cable designed to be buried.

14.10.4 Coaxial Cable Termination Techniques

14.10.4.1 Soldering

Soldering offers several advantages as it can be used on solid or stranded conductors and it creates both a solid mechanical and electrical connection. The disadvantage is that it takes more time to terminate than other methods and cold solder joints can cause problems if the connector is not soldered to the cable properly. The use of lead-based solder might also be a consideration if RoHS (reduction of hazardous substances) requirements are part of the installation. Soldering is not recommended for high-frequency applications, such as HD-SDI or 1080p/60 as the variations in dimensions will show up as variations in impedance and contribute to return loss (see Section 14.10.5).

14.10.4.2 Crimping

Crimping is probably the most popular method of terminating BNC and F connectors on coax cable. Like the solder method, it can be used on solid or stranded conductors and provides a good mechanical and electrical connection. This method is the most popular because there is no need for soldering so installation time is reduced. It is very important to use the proper size connector for a tight fit on the cable. Always use the proper tool. Never use pliers as they are not designed to place the pressure of the crimp evenly around the connector. Pliers will crush the cable and can degrade the electrical properties of the cable.

14.10.4.3 Twist-On Connectors

Twist-on connectors are the quickest way of terminating a coaxial cable; however, they do have some drawbacks. When terminating the cable with this type of connector, the center conductor is scored by the center pin on the connector, thus too much twisting can cause damage to the center conductor. It is not recommended for pan and tilt installations as the constant movement of the cable may work the connector loose. Because there is no mechanical or electrical crimp or solder connection, this connector is not as reliable as the other methods.

14.10.4.4 Compression Connectors

There are connectors, often a one-piece connector, that fit over the stripped cable and fasten by having two parts squeeze or compress together. This is a very simple and reliable way of connecting cable. However, the very high-frequency performance (beyond 500 MHz) has yet to be proven and so these connectors are not recommended for professional digital applications. A compression connector that is measured with a return loss of –20 dB at 2 GHz would be acceptable for professional broadcast HD applications.

14.10.5 Return Loss

At high frequencies, where cable and connectors are a significant percentage of a wavelength, the impedance variation of cable and components can be a significant source of signal loss. When the signal sees something other than 75 Ω, a portion of the signal is reflected back to the source. Table 14-22 shows the wavelength and quarter wavelength at various frequencies. One can see that this was a minor problem with analog video (quarter wave 59 ft) since the distances are so long. However, with HD-SDI and higher signals, the quarter wave can be 1 inch or less, meaning that everything in the line is critical: cable connectors, patch panels, patch cords, adaptors, bulkhead/feedthrough connectors, etc.

Table 14-22. Wavelength and Quarter Wavelength of Various Signals at Various Frequencies

Signal	Clock Frequency	Third Harmonic	Wavelength	Quarter Wavelength
Analog video (4.2 MHz)	analog	analog	234 ft	59 ft

Table 14-22. Wavelength and Quarter Wavelength of Various Signals at Various Frequencies

Signal	Clock Frequency	Third Harmonic	Wavelength	Quarter Wavelength
SD-SDI	135 MHz	405 MHz	2.43 ft	7.3 inches
HD-SDI	750 MHz	2.25 GHz	5.3 inches	1.3 inches
1080p/60	1.5 GHz	4.5 GHz	2.6 inches	0.66 inches

In fact, Table 14-22 above is not entirely accurate. The distances should be multiplied by the velocity of propagation of the cable or other component, to get the actual length, so they are even shorter still.

Since everything is critical at high frequencies, it is appropriate to ask the manufacturers of the cable, connectors, patch panels, and other passive components, how close to 75 Ω their products are. This can be established by asking for the return loss of each component. Table 14-23 will allow the user to roughly translate the answers given.

Table 14-23. Return Loss versus % of Signal Received and Reflected

Return Loss	% of Signal Received	% Reflected
–50 dB	99.999%	0.001%
–40 dB	99.99%	0.01%
–30 dB	99.9%	0.1%
–20 dB	99.0%	1.0%
–10 dB	90.0%	10.0%

Most components intended for HD can pass –20 dB return loss. In fact, –20 dB return loss at 2 GHz is a good starting point for passive components intended for HD-SDI. Better components will pass –30 dB at 2 GHz. Better still (and rarer still) would be –30 dB at 3 GHz. There are currently no components that are consistently –40 dB return loss at any reasonable frequency. In Table 14-22, it can be seen that 1080p/60 signals need to be tested to 4.5 GHz. This requires expensive custom-built matching networks. As of this writing, only one company (Belden) has made such an investment.

Note that the number of nines in the Signal Received column is the same as the first digit of the return loss (i.e., –30 dB = 3 nines = 99.9%). There are similar tests, such as SRL (structural return loss). This test only partially shows total reflection. Do not accept values measured in any way except return loss. The SMPTE maximum amount of reflection on a passive line (with all components measured and added together) is –15 dB or 96.84% received, 3.16% reflected. A line with an RL of –10 dB (10% reflected) will probably fail.

14.10.6 Video Triaxial Cable

Video triaxial cable is used to interconnect video cameras to their related equipment. Triaxial cable contains a center conductor and two isolated shields, allowing it to support many functions on the one cable. The center conductor and outer shield carry the video signals plus intercoms, monitoring devices, and camera power. The center shield carries the video signal ground or common. Triax cable is usually of the RG-59 or RG-11 type.

14.10.7 S-Video

S-video requires a duplex (dual) coaxial cable to allow separate transmission of the luminance (Y) and the chrominance (C). The luminance signal is black or white or any gray value while the chrominance signal contains color information. This transmission is sometimes referred to as Y-C. Separating signals provides greater picture detail and resolution and less noise interference.

S-video is sometimes referred to as S-VHS™ (Super-Video Home System). While its intention was for improved consumer video quality, these cameras were also used for the lower end of the professional area, where they were used for news, documentaries, and other less-critical applications.

14.10.8 RGB

RGB stands for red-green-blue, the primary colors in color television. It is often called component video since the signal is split up to its component colors. When these analog signals are carried separately much better image resolution can be achieved. RGB can be carried on multiple single video cables, or in bundles of cables made for this application. With separate cables, all the cables used must be precisely the same electrical length. This may or may not be the same as the physical length. Using a vectorscope, it is possibly to determine the electrical length and compare the RGB components. If the cables are made with poor quality control, the electrical length of the coaxes may be significantly different (i.e., one cable may have to be physically longer than the others to align the component signals). Cables made with very good quality control can simply be cut at the same physical length.

Bundles of RGB cables should be specified by the amount of timing error, the difference in the delivery time on the component parts. For instance, all Belden bundled coax cables are guaranteed to be 5 ns (nanosecond) difference per 100 ft of cable. Other manufacturers should have a similar specification and/or guarantee. The de facto timing requirement for broadcast RGB is a maximum of 40 ns. Timing cables by hand with a vectorscope allows the installer to achieve timing errors of >1 ns. Bundled cables made for digital video can also be used for RGB analog, and similar signals (Y, R–Y, B–Y or Y, Pb, Pr or YUV or VGA, SVGA, XGA, etc.) although the timing requirements for VGA and that family of signals has not been established.

These bundled coaxes come in other version besides just three coax RGB. Often the horizontal and vertical synchronizing signals (H and V) are carried with the green video signal on the green coax. For even greater control, these signals can be carried by a single coax (often called RGBS) or five coaxes, one for each signal (called RGBHV). These cables are becoming more common in the home, where they are often referred to as five-wire video. There are also four-pair UTP data cables made especially to run RGB and VGA signals. Some of these have timing tolerance (called delay skew in the UTP world) that is seriously superior to bundled coaxes. However, the video signals would have to be converted from 75 Ω to 100 Ω, and the baluns to do this, one for each end of the cable, would be added to the cost of the installation. Further, the impedance tolerance of coax, even poorly made coax, is dramatically superior to twisted pairs. Even bonded twisted pairs are, at best, ±7 Ω, where most coaxial cables are ±3 Ω, with precision cables being twice as good as that, or even better.

14.10.9 VGA and Family

VGA stands for video graphics array. It is an analog format to connect progressive video source to displays, such as projectors and screens. VGA comes in a number of formats, based on resolution. These are shown in Table 14-24.

There are many more variations in resolution and bandwidth than the ones shown in Table 14-24.

14.11 Digital Video

There are many formats for digital video, for both consumer, commercial and professional applications. This section concentrates on the professional applications, mainly SD-SDI (standard definition—serial digital interface) and HD-SDI (high-definition—serial digital interface.) There are sections on related consumer standards such as DVI (Section 14.9.4.1) and HDMI (Section 14.9.4.2).

Table 14-24. Resolution of Various VGA and Family Formats

Signal Type	Resolution
VGA	640 × 480
SVGA	800 × 600
XGA	1024 × 768
WXGA	1280 × 720
SXGA	1280 × 1024
SXGA-HD	1600 × 1200
WSXGA	1680 × 1050
QXGA	2048 × 1536
QUSXG	3840 × 2400

14.11.1 Digital Signals and Digital Cable

Control communications, or data communications, uses digital signals. Digital video signals require wide bandwidth cabling. Control communications and data communications use lower-performance cabling because they carry less information, requiring less bandwidth. High-speed data communications systems have significant overhead added to handle error correction so if data is lost, it can be re-sent. Digital video has some error correction capabilities, however, if all of the data bits required to make the system work are not received, picture quality is reduced or lost completely. Table 14-25 compares various digital formats.

14.11.2 Coax and SDI

Most professional broadcast formats (SDI and HD-SDI) are in a serial format and use a single coaxial cable with BNC connectors. Emerging higher resolution formats, such as 1080p/60, are also BNC based. Some work with smaller connectors for dense applications, such as patch panels and routers, which use subminiature connectors such as LCC, DIN 1.0/2.3 or DIN 1.0/2.5. Proprietary miniature BNC connectors are also available.

14.11.3 Cables and SDI

The most common form of SDI, component SDI, operates at data rates of 270 Mbps (clock 135 MHz). Cable loss specifications for standard SDI are specified in SMPTE 259M and ITUR BT.601. The maximum cable length is specified as 30 dB signal loss at one-half the

Table 14-25. Comparing Coaxial Digital Formats

Standard	Format	Intended Use	Connector Style	Cable Type	Transmission Distance[2]	Sample Rate	Data Rate (Mbps)	Guiding Document
SDI	serial	broadcast	one BNC	coax[1]	300 m/1000 ft	27 MHz	270	SMPTE 259
SDTI	serial	data transport	one BNC	coax[1]	300 m/1000 ft	variable	270 or 360	SMPTE 305
SDTV	serial	broadcast	one BNC	coax[1]	300 m/1000 ft	27 MHz	3 to 8	ATSC; N53
HDTV	serial	broadcast	one BNC	coax[1]	122 m/400 ft	74.25 MHz	19.4	ATSC; A/53
HD-SDI	serial	broadcast	one BNC	coax[1]	122 m/400 ft	74.25 MHz	1500	SMPTE 292M
1080p/60	serial	Master format	one BNC	coax[1]	80 m/250 ft	148.5 MHz	3000	SMPTE 424M

[1] Also implemented over fiber systems

[2] Transmission distances may vary widely depending on cabling and the specific equipment involved.

clock frequency and is acceptable because serial digital receivers have signal recovery processing.

HD-SDI, whose cable loss is governed by SMPTE 292M, operates at a data rate of 1.5 Gbps (clock 750 MHz). The maximum cable length is specified at 20 dB signal loss at one-half the clock frequency. These are Manchester Coded signals and the bit rate is therefore double the clock rate. Emerging 1080p/60 applications are covered under SMPTE 424M. The data rate is 3 Gbps (clock 1.5 GHz).

14.11.4 *Receiver Quality*

The quality of the receiver is important in the final performance of a serial digital system. The receiver has a greater ability to equalize and recover the signal with SDI signals. SMPTE 292M describes the minimum capabilities of a type A receiver and a type B receiver. SDI receivers are considered adaptive because of their ability to amplify, equalize, and filter the information. Rise time is significantly affected by distance, and all quality receivers can recover the signal from a run of HD-SDI RG-6 (such as Belden 1694A) for a minimum distance of 122 m (400 ft). The most important losses that affect serial digital are rise time/fall time degradation and signal jitter. Serial digital signals normally undergo reshaping and reclocking as they pass through major network hubs or matrix routers.

Table 14-26 gives the specifications mandated in SMPTE 259M and SMPTE 292M in terms of rise/fall time performance and jitter. If the system provides this level of performance at the end of the cable run, the SDI receiver should be able to decode the signal.

swept at 2.25 GHz. RL can be no greater than 15 dB at

14.11.5 *Serial Digital Video*

Serial digital video (SDI) falls under standards by the Society of Motion Picture and Television Engineers (SMPTE) and ITU and falls under the following categories:

SMPTE 259M	Digital video transmissions of composite NTSC 143 Mb/s (Level A) and PAL 177 Mb/s (Level B). It also covers 525/625 component transmissions of 270 Mb/s (Level C) and 360 Mb/s (Level D).
SMPTE 292M	HDTV transmissions at 1.485 Gb/s
SMPTE 344M	Component widescreen transmission of 540 Mb/s
ITU-R BT.601	International standard for PAL transmissions of 177 Mb/s

These standards can work with standard analog video coax cables, however, the newer digital cables provide the more precise electrical characteristics required for high-frequency transmission.

SDI cable utilizes a solid bare-copper center-conductor which improves impedance stability and reduced return loss (RL). Digital transmissions contain both low-frequency and high-frequency signals so it is imperative that a solid-copper center-conductor is used rather than a copper-clad steel center conductor. This allows the low frequencies to travel down the center of the conductor and the high frequencies to travel on the outside of the conductor due to the skin effect. Since digital video consists of both low and high frequencies, foil shields work best. All SDI cable should be sweep tested for return loss to the third harmonic of the fundamental frequency. For HD-SDI which is 1.485 Gb/s or has a 750 MHZ bandwidth, the cable is this frequency.

Table 14-26. SMPTE Serial Digital Performance Specifications

	SMPTE 259				SMPTE 292M	
	Level A	Level B	Level C	Level D	Level D	Level L
Parameter	NTSC 4fsc Composite	PAL 4fsc Composite	525/625 Component	525/625 Component	1920 × 1080 Interlaced	1280 × 720 Progressive
Data Rate in Mbps (clock)	143	177	270	360	1485	1485
½ Clock Rate in MHz	71.5	88.5	135	180	742.5	742.5
Signal Amplitude (p-p)	800 mV	800 mV	800 mV	800 mV	800 mV	800 mV
dc Offset (volts)	0 ±0.5	0 ±0.5	0 ±0.5	0 ±0.5	0 ±0.5	0 ±0.5
Rise/Fall Time Max. (ns)	1.50	1.50	1.50	1.5	0.27	0.27
Rise/Fall Time Min. (ns)	0.40	0.40	0.40	0.40	–	–
Rise/Fall Time Differential (ns)	0.5	0.5	0.5	0.5	0.10	0.10
% Overshoot Max.	10	10	10	10	10	10
Timing Jitter (ns)	1.40	1.13	0.74	0.56	0.67	0.67
Alignment Jitter (ns)	1.40	1.13	0.74	0.56	0.13	0.13

BNC 50 Ω connectors are often used to terminate digital video lines. This is probably acceptable if only one or two connectors are used. However, if more connectors are used, 75 Ω connectors are required to eliminate RL. Connectors should exhibit a stable 75 Ω impedance out to 2.25 GHz, the third harmonic of 750 MHz.

14.12 Radio Guide Designations

From the late 1930s the U.S. Army and Navy began to classify different cables by their constructions. Since the intent of these high-frequency cables, both coaxes and twisted pairs, was to guide radio frequency signals, they carried the designation RG for radio guide.

There is no correlation between the number assigned and any construction factor of the cable. Thus an RG-8 came after an RG-7 and before an RG-9, but could be completely different and unrelated designs. For all intents and purposes, the number simply represents the page number in a book of designs. The point was to get a specific cable design, with predictable performance, when ordered for military applications.

As cable designs changed, with new materials and manufacturing techniques, variations on the original RG designs began to be manufactured. Some of these were specific targeted improvement, such as a special jacket on an existing design. These variations are noted by an additional letter on the designation. Thus RG-58C would be the third variant on the design of RG-58.

The test procedure for many of these military cables is often long, complicated, and expensive. For the commercial user of these cables, this is a needless expense. So many manufacturers began to make cables that were identical to the original RG specification except for testing. These were then designated utility grade and a slash plus the letter U is placed at the end. RG-58C/U is the utility version of RG-58C, identical in construction but not in testing.

Often the word *type* is included in the RG designation. This indicates that the cable under consideration is based on one of the earlier military standards but differs from the original design in some significant way. At this point, all the designation is telling the installer is that the cable falls into a family of cables. It might indicate the size of the center conductor, the impedance, and some aspects of construction, with the key word being *might*.

By the time the RG system approached RG-500, with blocks of numbers abandoned in earlier designs, the system became so unwieldy and unworkable that the military abandoned it in the 1970s and instituted MIL-C-17 (Army) and JAN C-17 (Navy) designations that continue to this day. RG-6, for instance, is found under MIL-C-17G.

14.13 Velocity of Propagation

Velocity of propagation, abbreviated V_p, is the ratio of the speed of transmission through the cable versus the speed of light in free space, about 186,282 miles per second (mi/s) or 299,792,458 meters per second (m/s). For simplicity, this is usually rounded up to 300,000,000 meters per second (m/s). Velocity of propagation is a good indication of the quality of the cable. Solid polyethylene has a V_p of 66%. Chemically formed foam has a

V_p of 78%, and nitrogen gas injected foam has a V_p up to 86%, with current manufacturing techniques. Some hardline, which is mostly dry air or nitrogen dielectric, can exceed 95% velocity.

Velocity of propagation is the velocity of the signal as it travels from one end of the line to the other end. It is caused because a transmission line, like all electrical circuits, possesses three inherent properties: resistance, inductance, and capacitance. All three of these properties will exist regardless of how the line is constructed. Lines cannot be constructed to eliminate these characteristics.

Under the foregoing conditions, the velocity of the electrical pulses applied to the line is slowed down in its transmission. The elements of the line are distributed evenly and are not localized or present in a lumped quantity.

The velocity of propagation (V_p) in flexible cables will vary from 50% to a V_p of 86%, depending on the insulating composition used and the frequency. V_p is directly related to the dielectric constant (DC) of the insulation chosen. The equation for determining the velocity of propagation is

$$V_p = \frac{100}{\sqrt{DC}} \tag{14-5}$$

where,

V_p is the velocity of propagation,

DC is the dielectric constant.

Velocity can apply to any cable, coax or twisted pairs, although it is much more common to be expressed for cables intended for high-frequency applications. The velocity of propagation of coaxial cables is the ratio of the dielectric constant of a vacuum to the square root of the dielectric constant of the insulator, and is expressed in percent.

$$\frac{V_L}{V_S} = \frac{1}{\sqrt{\varepsilon}} \tag{14-6}$$

or

$$V_L = \frac{V_S}{\sqrt{\varepsilon}} \tag{14-7}$$

where,

V_L is the velocity of propagation in the transmission line,

V_S is the velocity of propagation in free space,

ε is the dielectric constant of the transmission line insulation.

Various dielectric constants (ε) are as follows:

Material	Dielectric Constant
Vacuum	1.00
Air	1.0167
Teflon	2.1
Polyethylene	2.25
Polypropylene	2.3
PVC	3.0 to 6.5

14.14 Shielding

From outdoor news gathering to studios and control rooms to sound reinforcement systems, the audio industry faces critical challenges from EM/RF interference (EMI and RFI). Shielding cable and twisting pairs insures signal integrity and provides confidence in audio and video transmissions, preventing downtime and maintaining sound and picture clarity.

Cables can be shielded or unshielded, except for coaxial cable which is, by definition, a precise constructions of a shielded single conductor. There are a number of shield constructions available. Here are the most common.

14.14.1 Serve or Spiral Shields

Serve or spiral shield are the simplest of all wire-based shields. The wire is simply wound around the inner portions of the cable. Spiral shields can be either single or double spirals. They are more flexible than braided shields and are easier to terminate. Since spiral shields are, in essence, coils of wire, they can exhibit inductive effects which make them ineffective at higher frequencies. Therefore, spiral/serve shields are relegated to low frequencies and are rarely used for frequencies above analog audio. Serve or spiral shields tend to open up when the cable is bent or flexed. So shield effectiveness is less than ideal, especially at high frequencies.

14.14.2 Double Serve Shields

Serve or spiral shields can be improved by adding a second layer. Most often, this is run at a 90° angle to the original spiral. This does improve coverage although the tendency to open up is not significantly improved and so this is still relegated to low-frequency or analog audio applications. This double serve or spiral construction is also called a *Reussen* shield (pronounced roy-sen).

14.14.3 French Braid™

The French Braid shield by Belden is an ultraflexible double spiral shield consisting of two spirals of bare or tinned copper conductors tied together with one weave. The shield provides the long flex life of spiral shields and greater flexibility than braided shields. It also has about 50% less microphonic and triboelectric noise. Because the two layers are woven along one axis, they cannot open up as dual spiral/serve constructions can. So French Braid shields are effective up to high frequencies, and are used up to the Gigahertz range of frequencies.

14.14.4 Braid

Braid shields provide superior structural integrity while maintaining good flexibility and flex life. These shields are ideal for minimizing low-frequency interference and have lower dc resistance than foil. Braid shields are effective at low frequencies, as well as RF ranges. Generally, the higher the braid coverage, the more effective the shield. The maximum coverage of a single braid shield is approximately 95%. The coverage of a dual braid shield can be as much as 98%. One hundred percent coverage with a braid is not physically possible.

14.14.5 Foil

Foil shields can be made of bare metal, such as a bare copper shield layer, but more common is an aluminum-polyester foil. Foil shields can offer 100% coverage. Some cables feature a loose polyester-foil layer. Other designs can bond the foil to either the core of the cable or to the inside of the jacket of the cable. Each of these presents challenges and opportunities.

The foil layer can either face out, or it can be reversed and face in. Since foil shields are too thin to be used as a connection point, a bare wire runs on the foil side of the shield. If the foil faces out, the drain wire must also be on the outside of the foil. If the foil layer faces in, then the drain wire must also be inside the foil, adjacent to the pair.

Unbonded foil can be easily removed after cutting or stripping. Many broadcasters prefer unbonded foil layers in coaxial cable to help prevent thin slices of foil that can short out BNC connectors. If the foil is bonded to the core, the stripping process must be much more accurate to prevent creating a thin slice of core-and-foil.

However, with F connectors, which are pushed onto the end of the coax, unbonded foil can bunch up and prevent correct seating of these connectors. This explains why virtually all coaxes for broadband/CATV applications have the foil bonded to the core—so F connectors easily slip on.

In shielded paired cables, such as analog or digital audio paired cables, the foil shield wraps around the pair. Once the jacket has been stripped off, the next step is to remove the foil shield. These cables are also available where the foil is bonded (glued) to the inside of the jacket. When the jacket is removed, the foil is also removed, dramatically speeding up the process.

A shorting fold technique is often used to maintain metal-to-metal contact for improved high-frequency performance. Without the shorting fold, a slot is created through which signals can leak. A isolation fold also helps prevent the shield of one pair contacting the shield of an adjacent pair in a multipair construction. Such contact significantly increases crosstalk between these pairs.

An improvement on the traditional shorting fold used by Belden employs the Z-Fold™, designed for use in multipair applications to reduce crosstalk, Fig. 14-14. The Z-Fold combines an isolation fold and a shorting fold. The shorting fold provides metal-to-metal contact while the isolation fold keeps shields from shorting to one another in multipair, individually shielded cables.

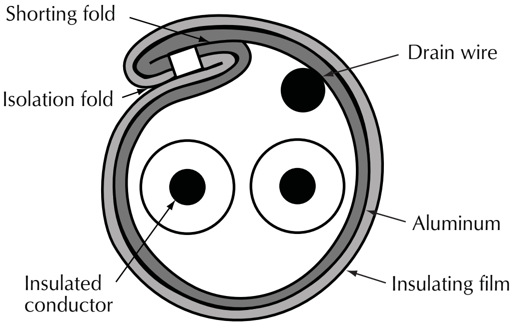


Figure 14-14. Z-Fold foil type shielded wire improves high frequency performance. Courtesy Belden.

Since the wavelength of high frequencies can eventually work through the holes in a braid, foil shields are most effective at those high frequencies. Essentially, foil shields represent a skin shield at high frequencies, where skin effect predominates.

14.14.6 Combination Shields

Combination shields consist of more than one layer of shielding. They provide maximum shield efficiency across the frequency spectrum. The combination foil-braid shield combines the advantages of 100% foil coverage and the strength and low dc resistance of a braid. Other combination shields available include various foil-braid-foil, braid-braid, and foil-spiral designs.

14.14.6.1 Foil + Serve

Because of the inductive effects of serve/spiral shields, which relegate them to low-frequency applications, this combination is rarely seen.

14.14.6.2 Foil + Braid

This is the most common combination shield. With a high-coverage braid (95%) this can be extremely effective over a wide range of frequencies, from 1 kHz to many GHz. This style is commonly seen on many cables, including precision video cable.

14.14.6.3 Foil + Braid + Foil

Foil-braid-foil is often called a tri-shield. It is most commonly seen in cable television (CATV) broadband coaxial applications. The dual layers of foil are especially effective at high frequencies. However, the coverage of the braid shield in between is the key to shield effectiveness. If it is a reasonably high coverage (>80%) this style of braid will have excellent shield effectiveness.

One other advantage of tri-shield coax cable is the ability to use standard dimension F connectors since the shield is essentially the same thickness as the common foil + braid shield of less expensive cables.

14.14.6.4 Foil + Braid + Foil + Braid

Foil-braid-foil-braid is often called *quad-shield* or just *quad* (not to be confused with starquad microphone cable or old POTS quad hookup cable). Like trishield above, this is most common in cable television (CATV) broadband coaxial applications. Many believe this to be the ultimate in shield effectives. However, this is often untrue.

If the two braids in this construction are high coverage braids (>80%) then, yes, this would be an exceptional cable. But most quad-shield cable uses two braids that are 40% and 60% coverage, respectively. With that construction, the tri-shield with an 80%+ braid is measurably superior. Further, quad-shield coaxial cables are considerably bigger in diameter and therefore require special connectors.

Table 14-27 shows the shield effectiveness of different shield constructions at various frequencies. Note that all the braids measured are aluminum braids except for the last cable mentioned. That last cable is a digital precision video (such as Belden 1694A) and is many times the cost of any of the other cables listed.

Table 14-27. Shield Effectiveness of Different Shield Constructions

Shield Type (Aluminum Braid)	5 MHz	10 MHz	50 MHz	100 MHz	500 MHz
60% braid, bonded foil	20	15	11	20	50
60% braid, tri-shield	3	2	0.8	2	12
60%/40% quad shield	2	0.8	0.2	0.3	10
77% braid, tri-shield	1	0.6	0.1	0.2	2
95% copper braid, foil	1	0.5	0.08	0.09	1

14.15 Shield Current Induced Noise

There is significant evidence that constructions that feature bonded foil with an internal drain wire may affect the performance of the pairs, especially at high frequencies. Since an ideal balanced line is one where the two wires are electrically identical, having a drain wire in proximity would certainly seem to affect the symmetry of the pair. This would be especially critical where strong RF fields are around audio cables.

Despite this evidence, there are very few cables made with appropriate symmetry. This may be based on lack of end-user demand, as manufacturers would be glad to redesign their cables should the demand arise. The drain wire could be easily substituted with a symmetrical low-coverage braid, for instance.

14.16 Grounds of Shields

With any combination shield, the braid portion is the part that is making the connection. Even if we are shielding against high-frequency noise, in which case the foil is doing the actual work, the noise gets to ground by way of the braid which is much lower in resistance than the foil.

Where the foil uses a drain wire, it is that drain wire that is the shield connection. Therefore, that drain wire must be bare so it can make contact with the foil. If the foil is floating, not glued or bonded to the core of the cable, then another plastic layer is used to carry the foil.

The foil itself is much too thin and weak to even be applied in the factory by itself. The second plastic layer adds enough strength and flex-life (flexes until failure) to allow the foil to be used.

The drain wire, therefore, must be in contact with the foil. In some cables, the foil faces out, so the drain wire must be on the outside of the foil, between the foil and the jacket. If the foil faces in, then the drain wire must be on the inside of the foil, adjacent to the pair (or other components) inside the cable.

With an internal drain wire, there are a number of additional considerations. One is SCIN, shield current induced noise, mentioned earlier in Section 14.8.5.1. Another is the ability to make a multipair shielded cable where the shields are facing in and the plastic facing out. This allows the manufacturer to color code the pairs by coloring the plastic holding the foil.

If you have a multipair cable, with individual foil shields, it is important that these foil shields do not touch. If the shields touch, then any signal or noise that is on one foil will be instantly shared by the other. You might as well put a foil shield in common around both pairs. Therefore, it is common to use foil shields facing in which will help prevent them from touching. These can then be color coded by using various colors of plastic with each foil to help identify each pair.

However, simply coiling the foil around the pair still leaves the very edge of the foil exposed. In a multipair cable with many individual foils, where the cable is bent and flexed to be installed, it would be quite easy for the edge of one foil to touch the edge of another foil, thus compromising shield effectiveness. The solution for this is a Z-fold invented by Belden in 1960, shown in Fig. 14-14. This does not allow any foil edge to be exposed no matter how the cable is flexed.

14.16.1 Ground Loops

In many installations, the ground potential between one rack and another, or between one point in a building and another, may be different. If the building can be installed with a star ground, the ground potential will be identical throughout the building. Then the connection of any two points will have no potential difference.

When two points are connected that do have a potential difference, this causes a ground loop. A ground loop is the flow of electricity down a ground wire from one point to another. Any RF or other interference on a rack or on an equipment chassis connected to ground will now flow down this ground wire, turning that foil or braid shield into an antenna and feeding that noise into the twisted pair. Instead of a small area of interference, such as where wires cross each other, a ground loop can use the entire length of the run to introduce noise.

If one cannot afford the time or cost of a star ground system, there are still two options. The first option is to cut the ground at one end of the cable. This is called a telescopic ground.

14.16.2 Telescopic Grounds

Where a cable has a ground point at each end, disconnecting one end produces a telescopic ground. Installers should be cautioned to disconnect only the destination (load) end of the cable, leaving the source end connected.

For audio applications, the effect of telescopic grounds will eliminate a ground loop, but at a 50% reduction in shield effectiveness (one wire now connected instead of two). If one disconnects the source end, which in analog audio is the low-impedance end, and maintains the destination (load) connection, this will produce a very effective R-L-C filter at audio frequencies.

At higher frequencies, such as data cables, even a source-only telescopic shield can have some serious problems. Fig. 14-15 shows the effect of a telescopic ground on a Cat 6 data cable. The left column shows the input impedance, the impedance presented to any RF traveling on the shield, at frequency F_k (bottom scale) in MHz.

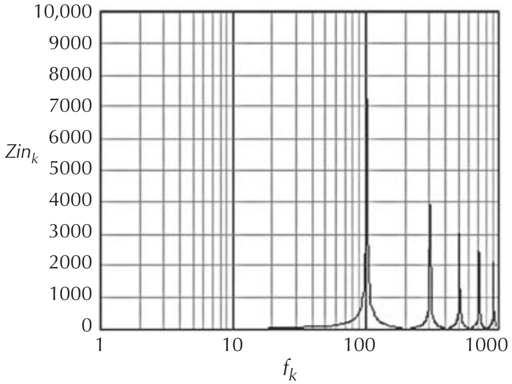


Figure 14-15. Effect of a telescopic ground on a Cat 6 cable.

You will note that at every half-wavelength, the shield acts like an open circuit. Since most audio cables are foil shielded, and the foil is effective only at high frequencies, this means that even a correctly terminated telescopic shield is less effective at RF frequencies.

14.17 UTP and Audio

One other solution for ground loops is to have no ground connection. For the seasoned audio and video professional, this solution may require a leap of faith. In can clearly be seen that, with a cable that has no shield, no drain wire, and no ground wire, so no ground loop can develop. This is a common form of data cable called UTP, unshielded twisted pairs.

With such a cable, having no shield means that you are totally dependent on the balanced line to reject noise. This is especially true, where you wish to use the four pairs in a Cat 5e, 6, or 6a cable to run four unrelated audio channels. Tests were performed on low-performance (stranded) Cat 5e patch cable (Belden 1752A) looking at crosstalk between the pairs. This test shows the average of all possible pair combinations, the worst possible case, and covered a bandwidth of 1 kHz to 50 kHz. The results are shown in Fig. 14-16.

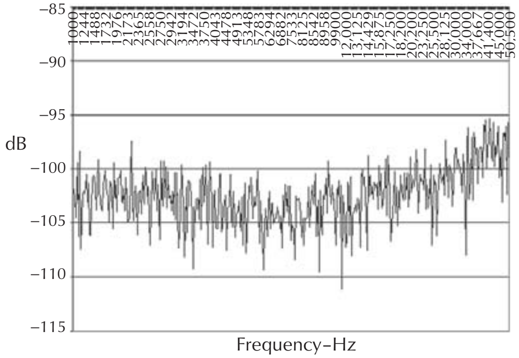


Figure 14-16. Crosstalk between Cat 5e patch cable.

You will note that the worst case is around 40 kHz where the crosstalk is slightly better than –95 dB. In the range of common audible frequencies (20 kHz) the pair-to-pair crosstalk approaches −100 dB. Since a noise floor of –90 dB is today considered wholly acceptable, a measurement of –95 dB or –100 dB is even better still.

A number of data engineers questioned these numbers based on the fact that these measurements were FEXT, far-end crosstalk, where the signals are weakest in such a cable. So measurements were also taken of NEXT, near-end crosstalk, where the signals are strongest. Those measurements are shown in Fig. 14-17.

The NEXT measurements are even better than the previous FEXT measurements. In this case, the worst case is exactly –95 dB at just under 50 kHz. At 20 kHz and below, the numbers are even better than the previous graph, around −100 dB or better.

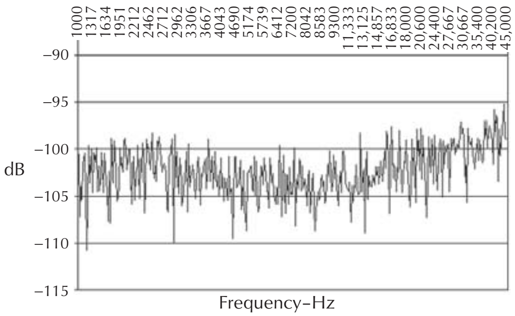


Figure 14-17. NEXT crosstalk.

There were attempts made to test a much better cable (Belden 1872A MediaTwist). This unshielded twisted-pair cable is now a Cat 6 bonded-pair design. After weeks of effort, it was determined that the pair-to-pair crosstalk could not be read on an Agilent 8714ES network analyzer. The crosstalk was somewhere below the noise floor of the test gear. The noise floor of that instrument is −110 dB. With a good cable, the crosstalk is somewhere below −110 dB.

14.17.1 So Why Shields?

These experiments with unshielded cable beg the question, why have a shield? In fact, the answer is somewhat reversed. The pairs in data cables are dramatically improved over the historic audio pairs. The bandwidth alone, 500 MHz for Cat 6a, for instance, indicates that these are not the same old pairs but something different. In fact, what has happened is that the wire and cable (and data) industries have fixed the pairs.

Before, with a poorly manufactured pair, a shield would help prevent signals from getting into, or leaking out of, a pair. The fact that either effect, ingress or egress, occurred indicated the poor balance, the poor performance of the pair.

This does not mean shields are dead. There are data cables with overall shields (FTP), even individually shielded pairs (Cat 7) common in Europe. However, these are subject to the same problems as all shielded, grounded cables in terms of ground loops and wavelength effects as shown in Sections 14.8.6.5 and 14.8.6.6.

The truth to the efficacy of unshielded twisted pairs running audio, video, data and many other signals is commonplace today. Many audio devices routinely use

UTP for analog and digital connections. Where the source is not a balanced line, a device must change from balanced (UTP) to unbalanced (coax, for instance). Such a device matches *Bal*anced-to-*Un*balanced and is therefore called a balun. There is more on baluns in Section 14.9.3.6.7, Baluns.

14.18 AES/EBU Digital Audio Cable

Digital audio technology has been around for many years, even decades, but until recently it has not been used much for audio. This has now changed and digital audio is overtaking analog audio. For this reason it is important that the cable used for digital signals meet the digital requirements. To set a standard, the Audio Engineering Society (AES) and the European Broadcast Union (EBU) have set standards for digital audio cable. The most common sampling rates and equivalent bandwidth are shows in Table 14-28.

It is important that the line impedance be maintained to eliminate reflections that degrade the signal beyond recovery. Standard analog cable can be used for runs under 50 ft (15 m) but beyond that, reliability decreases. The impedance and capacitance of analog cable is 40 to 70 Ω?and 20 to 50 pF/ft. The impedance and capacitance for digital cable is 110 Ω and 13 pF/ft with a velocity of propagation of 78%. Proper impedance match and low capacitance are required so the square wave signal is not distorted, reflected, or attenuated.

Broadcast cable is most often #24 (7 × 32) tinned copper wire with short overall twist lengths, low-loss foam insulation, and 100% aluminum polyester foil shield for permanent installations. Braided shields are also available for portable use. If required, #22 to #26 wire can be obtained. Digital audio cable also comes in multiple pairs with each pair individually shielded, and often jacketed, allowing each pair and its shield to be completely isolated from the others. One pair is capable of carrying two channels of digital audio. Cables are terminated with either XLR connectors or are punched down or soldered in patch panels.

14.18.1 AES/EBU Digital Coaxial Cable

Digital audio requires a much wider bandwidth than analog. As the sampling rate doubles, the bandwidth also doubles, as shown in Table 14-28.

Table 14-28. Sampling Rate versus Bandwidth

Sampling Rate kHz	Bandwidth MHz	Sampling Rate kHz	Bandwidth MHz
32.0	4.096	48.0	6.144
38.0	4.864	96.0	12.228
44.1	5.6448	192.0	24.576

Digital audio can be transmitted farther distances over coax than over twisted pairs. The coax should have a 75 Ω impedance, a solid copper center conductor, and have at least 90% shield coverage. When transmitting audio over an unbalanced coax line, the use of baluns may be required to change from balanced to unbalanced and back unless the device contains AES/EBU unbalanced coax inputs and outputs. The baluns change the impedance from 110 Ω balanced to 75 Ω unbalanced and back.

14.19 Triboelectric Noise

Noise comes in a variety of types such as EMI (electromagnetic interference) and RFI (radio frequency interference). There are also other kinds of noise problems that concern cables. These are mechanically generated or mechanically induced noise, commonly called *triboelectric noise*.

Triboelectric noise is generated by mechanical motion of a cable causing the wires inside the shield to rub against each other. Triboelectric noise is actually small electrical discharges created when conductors position changes relative to each other. This movement sets up tiny capacitive changes that eventually pop. Highly amplified audio can pick this up.

Fillers, nonconductive elements placed around the conductors, help keep the conductor spacing constant while semiconductive materials, such as carbon-impregnated cloth or carbon-plastic layers, help dissipate charge buildup. Triboelectric noise is measured through low noise test equipment using three low noise standards: NBS, ISA-S, and MIL-C-17.

Mechanically induced noise is a critical and frequent concern in the use of high-impedance cables such as guitar cords and unbalanced microphone cables that are constantly moving. The properties of special conductive tapes and insulations are often employed to help prevent mechanically induced noise. Cable without fillers can often produce triboelectric noise. This is why premise/data category cables are not suitable for flexing, moving audio applications. There are emerging flexible tactical data cables, especially those using bonded pairs, that might be considered for these applications.

14.20 Conduit Fill

To find the conduit size required for any cable, or group of cables, do the following:

1. Square the OD (outside diameter) of each cable and total the results.
2. To install only one cable: multiply that number by 0.5927.
3. To install two cables: multiply by 1.0134.
4. To install three or more cables: multiply the total by 0.7854.
5. From step #2 or #3 or #4, select the conduit size with an area equal to or greater than the total area. Use the ID (inside diameter) of the conduit for this determination.

This is based on the NEC ratings of

• single cable	53% fill
• two cables	31% fill
• three or more cables	40% fill

If the conduit run is 50 ft to 100 ft, reduce the number of cables by 15%. For each 90° bend, reduce the conduit length by 30 ft. Any run over 100 ft requires a pull box at some midpoint.

14.21 Long Line Audio Twisted Pairs

As can be seen in Table 14-29, low frequency signals, such as audio, rarely go a quarter-wavelength and, therefore, the attributes of a transmission line, such as the determination of the impedance and the loading/matching of that line, are not considered.

However, long twisted pairs are common for telephone and similar applications, and now apply for moderate data rate, such as DSL. A twisted-pair transmission line is loaded at stated intervals by connecting an inductance in series with the line. Two types of loading are in general usage—lumped and continuous. Loading a line increases the impedance of the line, thereby decreasing the series loss because of the conductor resistance.

Although loading decreases the attenuation and distortion and permits a more uniform frequency characteristic, it increases the shunt losses caused by leakage. Loading also causes the line to have a cutoff frequency above which the loss becomes excessive. In a continuously loaded line, loading is obtained by wrapping the complete cable with a high-permeability magnetic tape or wire. The inductance is distributed evenly along the line, causing it to behave as a line with distributed constants.

In the lumped loading method, toroidal wound coils are placed at equally spaced intervals along the line, as shown in Fig. 14-18. Each coil has an inductance on the order of 88 mH. The insulation between the line conductors and ground must be extremely good if the coils are to function properly.

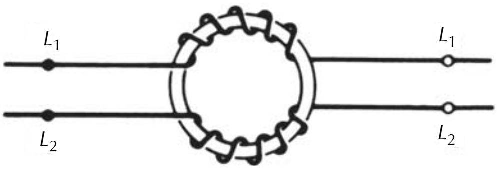


Figure 14-18. Loading coil connected in a balanced transmission line.

Loading coils will increase the talking distance by 35 to 90 miles for the average telephone line.

If a high-frequency cable is not properly terminated, some of the transmitted signal will be reflected back toward the transmitter, reducing the output.

14.22 Delay and Delay Skew

The fact that every cable has a velocity of propagation, obviously means that it takes time for a signal to go down a cable. That time is called *delay*, normally measured in nanoseconds (*Dn*). V_p can easily be converted into delay. Since V_p is directly related to dielectric constant (*DC*), they are all directly related as shown in Eq. 14-8 and determine the delay in nanoseconds- per-foot (ns/ft).

$$\begin{aligned} Dn &= \frac{100}{V_p} \\ &= \sqrt{DC} \end{aligned} \tag{14-8}$$

While these equations will give you a reasonable approximate value, the actual equations should be

$$\begin{aligned} Delay &= \frac{101.67164}{V_p} \\ &= 1.0167164\sqrt{DC}. \end{aligned} \tag{14-9}$$

Delay becomes a factor in broadcasting when multiple cables carry a single signal. This commonly occurs in RGB or other component video delivery systems. Delay also appears in high-data rate UTP, such a 1000Base-T (1GBase-T) and beyond where data is split between the four pairs and combined at the destination device.

Where signals are split up and recombined, the different cables supplying the components will each have a measurable delay. The trick is for all the component cables to have the *same delay* to deliver their portions at the same time. The de facto maximum timing variation in delay for RGB analog is delivery of all components within 40 ns. Measuring and adjusting cable delivery is often called timing. By coincidence, the maximum delay difference in the data world is 45 ns, amazingly close. In the data world, this is called *skew* or *delay skew*, where delivery does not line up.

In the RGB world, where separate coax cables are used, they have to be cut to the same electrical length. This is not necessarily the same physical length. Most often, the individual cables are compared by a Vectorscope, which can show the relationship between components, or a TDR (time domain reflectometer) that can establish the electrical length (delay) of any cable.

Any difference in physical versus electrical length can be accounted for by the velocity of propagation of the individual coaxes, and therefore, the consistency of manufacture. If the manufacturing consistency is excellent, then the velocity of all coaxes would be the same, and the physical length would be the same as the electrical length. Where cables are purchased with different color jackets, to easily identify the components, they are obviously made at different times in the factory. It is then a real test of quality and consistency to see how close the electrical length matches the physical length.

Where cables are bundled together, the installer then has a much more difficult time in reducing any timing errors. Certainly in UTP data cables, there is no way to adjust the length of any particular pair. In all these bundled cables, the installer must cut and connectorize.

This becomes a consideration when four-pair UTP data cables (category cables) are used to deliver RGB, VGA, and other nondata component delivery systems. The distance possible on these cables is therefore based on the attenuation of the cables at the frequency of operation, and on the delay skew of the pairs. Therefore, the manufacturers measurement and guarantee (if any) of delay skew should be sought if nondata component delivery is the intended application.

14.23 Attenuation

All cable has attenuation and the attenuation varies with frequency. Attenuation can be found with the equation

$$A = 4.35\frac{R_t}{Z_o} + 2.78pf\sqrt{\varepsilon} \qquad (14\text{-}10)$$

where,

A is the attenuation in dB/100 ft,

R_t is the total dc line resistance in Ω/100 ft,

ε is the dielectric constant of the transmission line insulation,

p is the power factor of the dielectric medium,

f is the frequency,

Z_o is the impedance of the cable.

Table 14-29 gives the attenuation for various 50 Ω, 52 Ω, and 75 Ω cables. The difference in attenuation is due to either the dielectric of the cable or center-conductor diameter.

14.24 Characteristic Impedance

The *characteristic impedance* of a cable is the measured impedance of a cable of infinite length. This impedance is an ac measurement, and cannot be measured with an ohmmeter. It is frequency-dependent, as can be seen in Fig. 14-19. This shows the impedance of a coaxial cable from 10 Hz to 100 MHz.

At low frequencies, where resistance is a major factor, the impedance is changing from a high value (approximately 4000 Ω at 10 Hz) down to a lower impedance. This is due to skin effect (see Section 14.2.8), where the signal is moving from the whole conductor at low frequencies to just the skin at high frequencies. Therefore, when only the skin is carrying the signal, the resistance of the conductor is of no importance. This can be clearly seen in the equations for impedance, Eq. 14-13, for low frequencies, shows R, the resistance, as a major component. For high frequencies, Eq. 14-14, there is no R, no resistance, even in the equation.

Once we enter that high-frequency area where resistance has no effect, around 100 kHz as shown in Fig. 14-19, we enter the area where the impedance will not change. This area is called the *characteristic impedance* of the cable.

The characteristic impedance of an infinitely long cable does not change if the far end is open or shorted. Of course, it would be impossible to test this as it is impossible to short something at infinity. It is important to terminate coaxial cable with its rated impedance or a portion of the signal can reflect back to the input, reducing the efficiency of the transmission. Reflections can be caused by an improper load, using a wrong connector—i.e., using a 50 Ω video BNC connector at

Table 14-29. Coaxial Cable Signal Loss (Attenuation) in dB/100 ft

Frequency	RG-174/ 8216	RG-58/ 8240	RG-8X/ 9258	RG-8/ 8237	RG-8/ RF-9913F	RG-59/ 8241	RG-6/ 9248	RG-11/ 9292
1 MHz	1.9	0.3	0.3	0.2	0.1	0.6	0.3	0.2
10 MHz	3.3	1.1	1.0	0.6	0.4	1.1	0.7	0.5
50 MHz	5.8	2.5	2.3	1.3	0.9	2.4	1.5	1.0
100 MHz	8.4	3.8	3.3	1.9	1.3	3.4	2.0	1.4
200 MHz	12.5	5.6	4.9	2.8	1.8	4.9	2.8	2.1
400 MHz	19.0	8.4	7.6	4.2	2.7	7.0	4.0	2.9
700 MHz	27.0	11.7	11.1	5.9	3.6	9.7	5.3	3.9
900 MHz	31.0	13.7	13.2	6.9	4.2	11.1	6.1	4.4
1000 MHz	34.0	14.5	14.3	7.4	4.5	12.0	6.5	4.7
Characteristic impedance—Ω	50.0	50.0	50.0	52.0	50.0	75.0	75.0	75.0
Velocity of propagation—%	66	66%	80%	66%	84%	66%	82%	78%
Capacitance pF/ft, pF/m	30.8/101.0	29.9/98.1	25.3/83.0	29.2/96.8	24.6/80.7	20.5/67.3	16.2/53.1	17.3/56.7

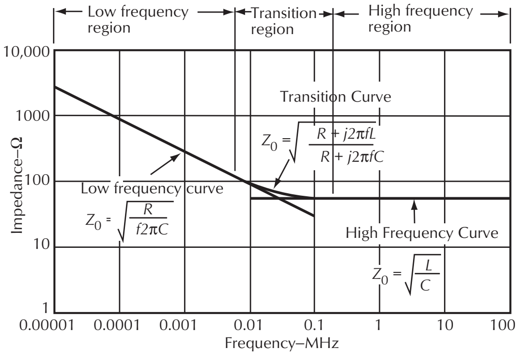


Figure 14-19. Impedance of coaxial cable from 10 Hz to 100 MHz.

high frequencies rather than a 75 Ω connector—a flattened cable, or too tight a bend radius, which changes the spacing between the conductors. Anything that affects the dimensions of the cable, will affect the impedance and create reflective losses. It would just be a question of how much reflection is caused. Reflections thus caused are termed *return loss*.

The characteristic impedance of common coaxial cable can be between 30 Ω and 200 Ω. The most common values are 50 Ω and 75 Ω. The characteristic Z_o is the average impedance of the cable equal to

$$Z_o = \frac{138}{\sqrt{\varepsilon}} \log\frac{D}{d} \tag{14-11}$$

where,

ε is the dielectric constant,

D is the diameter of the inner surface of the outer coaxial conductor (shield) in inches,

d is the diameter of the center conductor in inches.

The true characteristic impedance, at any frequency, of a coaxial cable is found with the equation

$$Z_o = \sqrt{\frac{R + j2\pi fL}{G + j2\rho C}} \tag{14-12}$$

where,

R is the series resistance of the conductor in ohms per unit length,

f is the frequency in hertz,

L is the inductance in henrys,

G is the shunt conductance in mhos per unit length,

C is the capacitance in farads.

At low frequencies, generally below 100 kHz, the equation for coaxial cable simplifies to

$$Z_o = \sqrt{\frac{R}{j2\rho\pi C}} \tag{14-13}$$

At high frequencies, generally above 100 kHz, the equation for coaxial cable simplifies to

$$Z_o = \sqrt{\frac{L}{C}} \tag{14-14}$$

14.25 Characteristic Impedance

The *characteristic impedance* of a transmission line is equal to the impedance that must be used to terminate the line in order to make the input impedance equal to the terminating impedance. For a line that is longer than a quarter-wavelength at the frequency of operation, the input impedance will equal the characteristic impedance of the line, irrespective of the terminating impedance.

This means that low-frequency applications often have quarter-wavelength distance way beyond common practical applications. Table 14-30 shows common signals, with the wavelength of that signal and the quarter-wavelength. To be accurate, given a specific cable type, these numbers would be multiplied by the velocity of propagation.

The question is very simple: will you be going as far as the quarter-wavelength, or farther? If so, then the characteristic impedance becomes important. As that distance gets shorter and shorter, this distance becomes critical. With smaller distances, patch cords, patch panels, and eventually the connectors themselves become just as critical as the cable. The impedance of these parts, especially when measured over the desired bandwidth, becomes a serious question. To be truly accurate, the quarter-wavelength numbers in Table 14-28 need to be multiplied by the velocity of propagation of each cable. So, in fact, the distances would be even shorter than what is shown.

Table 14-30. Characteristics of Various Signals

Signal Type	Bandwidth	Wave-length	Quarter-Wave-length	Quarter-Wave-length
Analog audio	20 kHz	15 km	3.75 km	12,300 ft
AES 3—44.1 kHz	5.6448 MHz	53.15 m	13.29 m	44 ft
AES 3—48 kHz	6.144 MHz	48.83 m	12.21 m	40 ft
AES 3—96 kHz	12.288 MHz	24.41 m	6.1 m	20 ft
AES 3—192 kHz	24.576 MHz	12.21 m	3.05 m	10 ft
Analog video (U.S.)	4.2 MHz	71.43 m	17.86 m	59 ft
Analog video (PAL)	5 MHz	60 m	15 m	49.2 ft
SD-SDI	135 MHz clock	2.22 m	55.5 cm	1 ft 10 in
SD-SDI	405 MHz third harmonic	74 cm	18.5 cm	7.28 in
HD-SDI	750 MHz clock	40 cm	10 cm	4 in
HD-SDI	2.25 GHz third harmonic	13 cm	3.25 cm	1.28 in
1080P/50-60	1.5 GHz clock	20 cm	5 cm	1.64 in
1080P/50-60	4.5 GHz third harmonic	66 mm	16.5 mm	0.65 in

It is quite possible that a cable can work fine with lower-bandwidth applications and fail when used for higher-frequency applications. The characteristic impedance will also depend on the parameters of the pair or coax cable at the applied frequency. The resistive component of the characteristic impedance is generally high at the low frequencies as compared to the reactive component, falling off with an increase of frequency, as shown in Fig. 14-19. The reactive component is high at the low frequencies and falls off as the frequency is increased.

The impedance of a uniform line is the impedance obtained for a long line (of infinite length). It is apparent, for a long line, the current in the line is little affected by the value of the terminating impedance at the far end of the line. If the line has an attenuation of 20 dB and the far end is short circuited, the characteristic impedance as measured at the sending end will not be affected by more than 2%.

14.26 Twisted-Pair Impedance

For shielded and unshielded twisted pairs, the characteristic impedance is

$$Z_0 = \frac{101670}{C(V_p)} \tag{14-15}$$

where,

Z_0 is the average impedance of the line,

C is found with Eqs. 14-16 and 14-17,

V_P is the velocity of propagation.

For unshielded pairs

$$C = \frac{3.68\varepsilon}{\log\left[\frac{2(ODi)}{DC(Fs)}\right]}. \tag{14-16}$$

For shielded pairs

$$C = \frac{3.68\varepsilon}{\log\left[\frac{1.06(ODi)}{DC(Fs)}\right]} \tag{14-17}$$

where,

ε is the dielectric constant,

ODi is the outside diameter of the insulation,

DC is the conductor diameter,

Fs is the conductor stranding factor (solid = 1, 7 strand = 0.939, 19 strand = 0.97.

The impedance for higher-frequency twisted-pair data cables is

$$Z_0 = 276\left(\frac{VP}{100}\right) \times \log\left[2\left(\frac{h}{DC \times Fs}\right) \times \left(\frac{1 - \left(\frac{h}{DC + Fb}\right)^2}{1 + \left(\frac{h}{DC + Fb}\right)^2}\right)\right] \tag{14-18}$$

where,

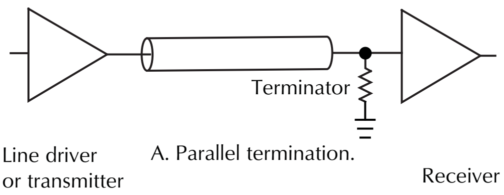
h is the center to center conductor spacing,

Fb is very near 0. Neglecting Fb will not introduce appreciable error.

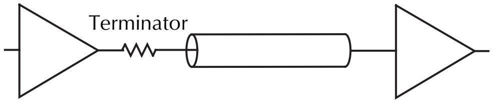
14.26.1 Transmission Line Termination

All lines do not need to be terminated. Knowing when to terminate a transmission line is a function of the frequency/wavelength of the signal and the length of the transmission line. Table 14-30 can be guideline, especially where the signal is long compared to the length of the line. If the wavelength of the signal is small compared to the transmission-line length, for instance a 4.5 GHz signal, a terminator is required to prevent the signal from reflecting back toward the source and interfering with forward traveling signals. In this case the line must be terminated for any line longer than a quarter of a wavelength.

Transmission-line termination is accomplished using parallel or series termination. Parallel termination connects a resistor between the transmission line and ground at the receiving end of the transmission line while series termination connects a resistor in series with the signal path near the beginning of the transmission line, Fig. 14-20.



A. Parallel termination.



B. Series termination.

Figure 14-20. Basic termination of transmission lines.

Resistive termination requires a resistor value that matches the characteristic impedance of the transmission line, most commonly a 50 Ω or 75 Ω characteristic impedance. The termination resistance is matched to the transmission line characteristic impedance so the electrical energy in the signal does not reflect back from the receiving end of the line to the source. If the resistor is perfectly matched to the characteristic impedance, at all frequencies within the desired bandwidth, all of the energy in the signal dissipates as heat in the termination resistor so no signal reflects backwards down the line to the source causing cancellations.

14.27 Loudspeaker Cable

Much has been said about wire for connecting loudspeakers to amplifiers. Impedance, inductance, capacitance, resistance, loading, matching, surface effects, etc. are constantly discussed.

Most home and studio loudspeaker runs are short (less than 50 ft, or 15 m) and therefore do not constitute a transmission line. When runs are longer, it is common to connect the loudspeakers as a 70 V, or 100 V system to reduce line loss caused by the wire resistance so the power lost in the line does not appreciably effect the power delivered to the loudspeaker. For instance, if a 4 Ω loudspeaker is connected to an amplifier with a cable which measures 4 Ω resistance, 50% of the power will be dissipated in the cable. If the loudspeaker was connected to a 70 V system, and the loudspeaker was taking 50 W from the amplifier, the loudspeaker/transformer impedance would be 100 Ω; therefore, the 4 Ω line resistance would dissipate 4% of the power.

When using a 70.7 V loudspeaker system, the choice of wire size for loudspeaker lines is determined by an economic balance of the cost of copper against the cost of power lost in the line. Power taken from the amplifier is calculated from the equation

$$P = \frac{V^2}{Z} \tag{14-19}$$

where,

P is the power delivered by the amplifier,

V is the voltage delivered by the amplifier,

Z is the impedance of the load.

For a 70 V system

$$P = \frac{5000}{Z} \tag{14-20}$$

If the voltage is 70.7 V and the load is 50 Ω, the power would be 100 W. However, if the amplifier was connected to a 50 Ω load with 1000 ft of #16 wire (2000 ft round trip) or 8 Ω of wire resistance the power from the amplifier would be

$$\begin{aligned} P &= \frac{5000}{50\Omega + 8\Omega} \\ &= 86.2\ \text{W} \end{aligned}$$

The current through the system is found with

$$I = \sqrt{\frac{P}{R}} \tag{14-21}$$

or in this case

$$\begin{aligned} I &= \sqrt{\frac{P}{R}} \\ &= \sqrt{\frac{86.2}{58}} \\ &= 1.21\ \text{A} \end{aligned}$$

The power to the 50 Ω load would be found with

$$P = I^2R \tag{14-22}$$

or in this case

$$\begin{aligned} P &= I^2R \\ &= 1.21^2R \\ &= 74.3W \end{aligned}$$

or 26% less power than assumed.

Only 11.7 W are lost to the line, the other 14 W cannot be taken from the amplifier because of the impedance mismatch. While high-power amplifiers are relatively inexpensive, it is still practical to use heavy enough wire so the amplifier can output almost its full power to the loudspeakers. Table 14-31 shows the characteristics of various cables which could be used for loudspeaker wire and Table 14-32 is a cable selection guide for loudspeaker cable.

Table 14-31. Frequency Limitations for 33 ft (10 m) Lengths of Cable with Various Loads

Cable Type	Upper Corner Frequency, kHz		Resonant Measurement Phase (°)	
	2 Ω Load	4 Ω Load	4 μF Load	4 Ω Load
No. 18 zip cord	75	136	35	3
No. 16 zip cord	61	114	32	2
No. 14 speaker cable	82	156	38	2
No. 12 speaker cable	88	169	40	2
No. 12 zip cord	55	106	32	4
Welding cable	100	200	44	2
Braided cable	360	680	80	1
Coaxial dual cylindrical	670	1300	112	
Coaxial RG-8	450	880	92	

Table 14-32. Loudspeaker Cable Selection Guide. Courtesy Belden.

Power (%)	11%	21%	50%
Loss (dB)	0.5	1.0	3.0
Wire Size	Maximum Cable Length–ft		
4 Ω Loudspeaker			
12 AWG	140	305	1150
14 AWG	90	195	740
16 AWG	60	125	470
18 AWG	40	90	340
20 AWG	25	50	195
22 AWG	15	35	135
24 AWG	10	25	85
8 Ω Loudspeaker			
12 AWG	285	610	2285
14 AWG	185	395	1480
16 AWG	115	250	935
18 AWG	85	190	685
20 AWG	50	105	390
22 AWG	35	70	275
24 AWG	20	45	170
70 V Loudspeaker			
12 AWG	6920	14,890	56,000
14 AWG	4490	9650	36,300
16 AWG	2840	6100	22,950
18 AWG	2070	4450	16,720
20 AWG	1170	2520	9500
22 AWG	820	1770	6650

The following explains how to use Table 14-29:

1. Select the appropriate speaker impedance column.
2. Select the appropriate power loss column deemed to be acceptable.
3. Select the applicable wire gage size and follow the row over to the columns determined in steps one and two. The number listed is the maximum cable run length.
4. The maximum run for 12 AWG in a 4 Ω loudspeaker system with 11% or 0.5 dB loss is 140 ft.

14.27.1 Damping Factor

The damping factor of an amplifier is the ratio of the load impedance (loudspeaker plus wire resistance) to the amplifier internal output impedance. The damping factor of the amplifier acts as a short circuit to the loudspeaker, controlling the overshoot of the loudspeaker. Present day amplifiers have an output impedance of less than 0.05 Ω which translates to a damping factor over 150 at 10 kHz, for instance, so they effectively dampen the loudspeaker as long as the loudspeaker is connected directly to the amplifier. Damping factor is an important consideration when installing home systems, studios, or any system where high-quality sound, especially at the low frequencies, is desired. As soon as wire resistance is added to the circuit, the damping factor reduces dramatically, reducing its affect on the loudspeaker. For instance, if a #16 AWG 50 ft loudspeaker cable (100 ft round trip) is used, the wire resistance would be 0.4 Ω making the damping factor only 18, considerably less than anticipated.

It is not too important to worry about the effect the damping factor of the amplifier has on the loudspeakers in a 70 V system as the 70 V loudspeaker transformers wipe out the effects of the wire resistance.

Consider the line as a lump sum, Fig. 14-21. The impedance of the line varies with wire size and type. Table 14-33 gives typical values of *R*, *C*, and *L* for 33 ft (10 m) long cables. Note, the impedance at 20 kHz is low for all but the smallest wire and the –3 dB upper frequency is well above the audio range. The worst condition is with a capacitive load. For instance, with a 4 µF load, resonance occurs around 35 kHz.

The results of the above are as follows:

1. Make the amplifier to loudspeaker runs as short as possible.
2. Use a wire gage that represents less than 5% of the loudspeaker impedance at any frequency.
3. Use twisted pairs on balanced 70 or 100 V distributed systems to reduce crosstalk (amplifier output is often fed back into the amplifier as negative feedback).
4. Use good connectors to reduce resistance.

Table 14-34 gives the length of cable run you can have for various loudspeaker impedances.

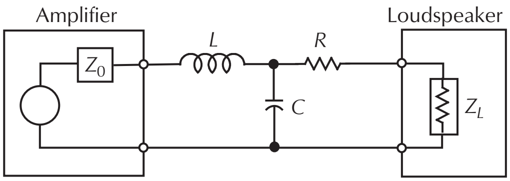


Figure 14-21. Amplifier, cable, loudspeaker circuit using lumped circuit elements to represent the properties of the cable

Table 14-33. Lumped Element Values for 33 ft (10 m) Lengths of Cable

Cable Type	*L*–µ*H*	*C*–*pF*	R_{dc}–Ω	*Z*–Ω@*20 kHz*
No. 18 zip cord	5.2	580	0.42	0.44
No. 16 zip cord	6.0	510	0.26	0.30
No. 14 speaker cable	4.3	570	0.16	0.21
No. 12 speaker cable	3.9	760	0.10	0.15
No. 12 zip cord	6.2	490	0.10	0.15
Welding cable	3.2	880	0.01	0.04
Braided cable	1.0	16,300	0.26	0.26
Coaxial dual cylindrical	0.5	58,000	0.10	0.10
Coaxial RG-8	0.8	300	0.13	0.13

14.27.2 Crosstalk

When a plurality of lines, carrying different programs or signals, are run together in the same conduit, or where multiple pairs or multiple coax cables are bundled, they tend to induce crosstalk currents into each other. Crosstalk is induced by two methods:

1. Electromagnetically through unbalanced coupling between one circuit and others.
2. Electrostatically through unbalanced capacitance to other circuits, or to the conduit if it carries current. This develops a voltage difference between one circuit and the others, or to its own or other shields carrying current.

If the line is less than a quarter-wavelength at the frequency of operation, then the cable does not have to have a specific impedance, or be terminated in a specific impedance. The terminating impedance could then be small compared to the open line characteristic impedance. The net coupling with unshielded pairs would then be predominantly magnetic. If the terminating impedance is much larger than the characteristic impedance of the wires, the net coupling will be predominantly electric.

Two wires of a pair must be twisted; this insures close spacing and aids in canceling pickup by transposition. In the measurements in Fig. 14-22, all pickup was

Table 14-34. Loudspeaker Cable Selection Guide

Power (%)	11%	21%	50%
Loss (dB)	0.5	1.0	3.0
Wire Size	Maximum Cable Length–ft		
4 Ω Loudspeaker			
12 AWG	140	305	1150
14 AWG	90	195	740
16 AWG	60	125	470
18 AWG	40	90	340
20 AWG	25	50	195
22 AWG	15	35	135
24 AWG	10	25	85
8 Ω Loudspeaker			
12 AWG	285	610	2285
14 AWG	185	395	1480
16 AWG	115	250	935
18 AWG	85	190	685
20 AWG	50	105	390
22 AWG	35	70	275
24 AWG	20	45	170
70 V Loudspeaker			
12 AWG	6920	14,890	56,000
14 AWG	4490	9650	36,300
16 AWG	2840	6100	22,950
18 AWG	2070	4450	16,720
20 AWG	1170	2520	9500
22 AWG	820	1770	6650

Courtesy Belden.

capacitive because the twisting of the leads effectively eliminated inductive coupling.

One application that is often ignored regarding crosstalk is speaker wiring, especially 70 V distributed loudspeaker wiring. You will note in the first drawing that the two wires are not a balanced line. One is hot the other is ground. Therefore, that pair would radiate some of the audio into the adjoining pair, also unbalanced. Twisting the pairs in this application would do little to reduce crosstalk.

The test was made on a 250 ft twisted pair run in the same conduit with a similar twisted pair, the latter carrying signals at 70.7 V. Measurements made for half this length produced half the voltages, therefore the results at 500 ft and 1000 ft were interpolated.

The disturbing line was driven from the 70 V terminals of a 40 W amplifier and the line was loaded at the far end with 125 Ω, thus transmitting 40 W. The crosstalk figures are for 1 kHz. The voltages at 100 Hz and 10 kHz are one-tenth and ten times these figures, respectively.

There are two ways to effectively reduce crosstalk. One is to run signals only on balanced-line twisted pairs. Even shielding has a small added advantage compared to the noise and crosstalk rejection of a balanced line. The second way to reduce crosstalk is to move the two cables apart. The inverse-square law tells us that doubling the distance will produce four times less interference. Further, if cables cross at right angles, this is the point where the magnetic fields have minimum interaction. Of course, the latter solution is not an option in a prebundled cable, or in cable trays or installations with multiple cables run from point to point.

14.28 National Electrical Code

The *National Electrical Code* (NEC) is a set of guidelines written to govern the installation of wiring and equipment in commercial buildings and residential areas. These guidelines were developed to insure the safety of humans as well as property against fires and electrical hazards. Anyone involved in specifying cable for installation should be aware of the basics of the code.

The NEC code book is made up of nine chapters, with each chapter divided into separate articles pertaining to specific subjects. Five articles pertain to communication and power-limited cable. The NEC book is written by and available from the NFPA (National Fire Protection Association), 11 Tracy Drive, Avon, MA 02322. They can be reached at 1-800-344-3555 or www.nfpa.org.

Article 725—Class 1, Class 2, Class 3, Remote-Control, Signaling, and Power-Limited Circuits. Article 725 covers Class 1, Class 2, and Class 3 remote control and signaling cables as well as power-limited tray cable. Power-limited tray cable can be used as a Class 2 or Class 3 cable. Cable listed multipurpose, communications, or power-limited fire protective can be used for Class 2 and Class 3 applications. A Class 3 listed cable can be used as a Class 2 cable.

Article 760—Fire Protective Signaling Systems. Article 760 covers power-limited fire-protective cable. Cable listed as power-limited fire-protective cable can also be used as Class 2 and Class 3 cable. Cable listed as communications and Class 3 can be used as power-limited fire protective cable with restrictions to conductor material and type gage size and number of conductors.

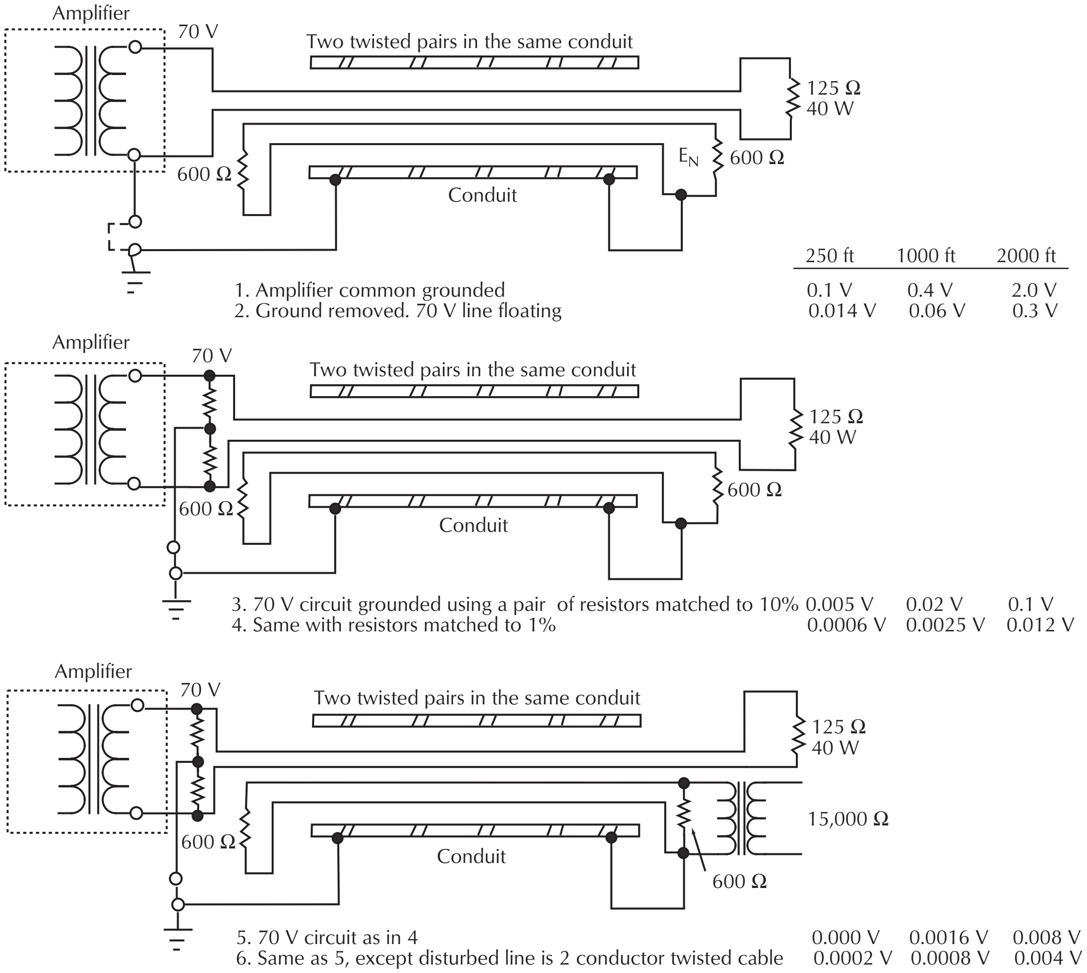

Amplifier
70 V
Two twisted pairs in the same conduit
125 Ω
40 W
E_N
600 Ω
600 Ω
Conduit
250 ft
1000 ft
2000 ft
1. Amplifier common grounded
0.1 V
0.4 V
2.0 V
2. Ground removed. 70 V line floating
0.014 V
0.06 V
0.3 V
Amplifier
70 V
Two twisted pairs in the same conduit
125 Ω
40 W
600 Ω
600 Ω
Conduit
3. 70 V circuit grounded using a pair of resistors matched to 10%
0.005 V
0.02 V
0.1 V
4. Same with resistors matched to 1%
0.0006 V
0.0025 V
0.012 V
Amplifier
70 V
Two twisted pairs in the same conduit
125 Ω
40 W
15,000 Ω
600 Ω
600 Ω
Conduit
5. 70 V circuit as in 4
0.000 V
0.0016 V
0.008 V
6. Same as 5, except disturbed line is 2 conductor twisted cable
0.0002 V
0.0008 V
0.004 V

Figure 14-22. Effects of grounding on crosstalk. Courtesy Altec Lansing Corp.

Article 770—Fiber Optic Systems. Article 770 covers three general types of fiber optic cable: nonconductive, conductive, and composite. Nonconductive type refers to cable containing no metallic members and no other electrically conductive materials. Conductive type refers to cable containing noncurrent carrying conductive members such as metallic strength members, etc. Composite type refers to cable containing optical fibers and current carrying electrical conductors. Composite types are classified according to the type of electrical circuit that the metallic conductor is designed for.

Article 800—Communication Circuits. Article 800 covers multipurpose and communication cable. Multipurpose cable is the highest listing for a cable and can be used for communication, Class 2, Class 3, and power-limited fire-protective cable. Communication cable can be used for Class 2 and Class 3 cable and also as a power-limited fire protective cable with restrictions.

Article 820—Community Antenna Television. Article 820 covers community antenna television and RF cable. CATV cable may be substituted with multipurpose or communication listed coaxial cable.

14.28.1 Designation and Environmental Areas

The NEC has designated four categories of cable for various environments and they are listed from the highest to the lowest listing. A higher listing can be used as a substitute for a lower listing.

Plenum—Suitable for use in air ducts, plenums, and other spaces used for environmental air without conduit and has adequate fire-resistant and low-smoke producing characteristics. It can also be substituted for all applications below.

Riser—Suitable for use in a vertical run, in a shaft, or from floor to floor, and has fire-resistant characteristics capable of preventing the spread of fire from floor to floor. It can also be substituted for all applications below.

General Purpose—Suitable for general-purpose use, with the exception of risers, ducts, plenums, and other space used for environmental air, and is resistant to the spread of fire. It can be substituted for the applications below.

Restricted Applications—Limited use and suitable for use in dwellings and in raceways and is flame retardant. Restricted use is limited to nonconcealed spaces of 10 ft or less, fully enclosed in conduit or raceway, or cable with diameters less than 0.25 inches for a residential dwelling.

14.28.2 Cable Types

Signal cable used for audio, telephone, video, control applications, and computer networks of less than 50 V is considered low-voltage cabling and is grouped into five basic categories by the NEC, Table 14-35.

Table 14-35. The Five Basic NEC Cable Groups

Cable Type	Use
CM	Communications
CL2, CL3	Class 2, Class 3 remote-control, signaling, and power-limited cables
FPL	Power-limited fire-protective signaling cables
MP	Multipurpose cable
PLTC	Power-limited tray cable

All computer network and telecommunications cabling falls into the CM class. The A/V industry primarily uses CM and CL2 cabling.

Table 14-36 defines the cable markings for various applications. Note plenum rated cable is the highest level because it has the lowest fire load which means it does not readily support fire.

14.28.3 NEC Substitution Chart

NEC cable hierarchy, Fig. 14-23 defines which cables can replace other cables. The chart starts with the highest listed cable on the top and descends to the lowest listed cable on the bottom. Following the arrows defines which cable can be substituted for others. Fig. 14-24 defines the Canadian Electrical Code (CEC) substitution chart.

Table 14-36. Cable Applications Designations Hierarchy

	Cable Family					
Application	MP	CM	CL2	CL3	FPL	PLTC
Plenum	MPP	CMP	CL2P	CL3P	FPLP	–
Riser	MPR	CPR	CL2R	CL3R	FPLR	–
General Purpose	MP, MPG	CM, CMG	CL2	CL3	FPL	PLTC
Dwelling	–	CMX	CL2X	CL3X	–	–

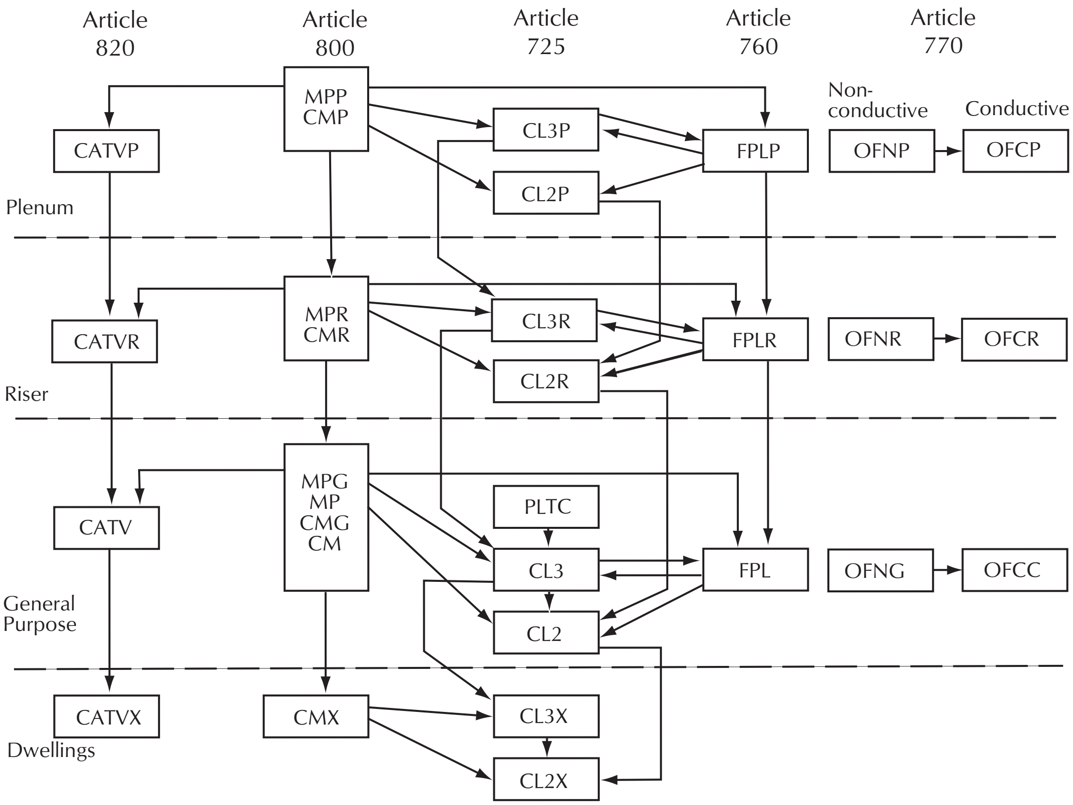
14.28.4 Final Considerations

The National Electrical Code is widely accepted as the suggested regulations governing the proper installation of wire and cable in the United States. The code is revised every three years to keep safety in the forefront in wire and cable manufacturing and installation. Even though the code is generally accepted, each state, county, city, and municipality has the option to adopt all of the code, part of the code, or develop one of its own. The local inspectors have final authority of the installation. Therefore, the NEC is a good reference when questions arise about the proper techniques for a particular installation, but local authorities should be contacted for verification.

When choosing cable for an installation, follow these three guidelines to keep problems to a minimum:

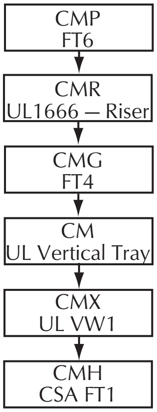
1. The application and environment determine which type of cable to use and what rating it should have. Make sure the cable meets the proper ratings for the application.
2. If substituting a cable with another, the cable must be one that is rated the same or higher than what the code calls for. Check with the local inspector as to what is allowed in the local area.
3. The NEC code is a general guideline that can be adopted in whole or in part. Local state, county, city, or municipal approved code is what must be followed. Contact local authorities for verification of the code in the area.

The local inspector or fire marshal has the final authority to approve or disapprove any installation of cable based on the National Electric Code or on the local code.


Article 820
Article 800
Article 725
Article 760
Article 770
Non-conductive
Conductive
MPP CMP
CATVP
CL3P
CL2P
FPLP
OFNP
OFCP
Plenum
MPR CMR
CATVR
CL3R
CL2R
FPLR
OFNR
OFCR
Riser
MPG MP CMG CM
CATV
PLTC
CL3
CL2
FPL
OFNG
OFCC
General Purpose
CATVX
CMX
CL3X
CL2X
Dwellings

MPP, MPR, MPG, MP—Multipurpose Cables
CMP, CMR, CMG, CM, CMX—Communications Cables
CL3P, CL3R, CL3, CL3X, CL2P, CL2R, CL2, CL2X—Class 2 and Class 3 Remote-Control, Signaling and Power Limited Cables
FPLP, FPLR, FPL—Power Limited Fire Alarm Cables
CATVP, CATVR, CATV, CATVX—Community Antenna Television and Radio Distribution Cables
OFNP, OFNR, OFNG, OFN—Nonconductive Optical Fiber Cables
OFCP, OFCR, OFCG, OFC—Conductive Optical Fiber Cables
PLTC—Power Limited Tray Cables

Figure 14-23. National Electrical Code substitution and hierarchy. Courtesy Belden.


CMP FT6
CMR UL1666 – Riser
CMG FT4
CM UL Vertical Tray
CMX UL VW1
CMH CSA FT1

The following cable substitutions may be used:

Communication cables marked MPP, CMP, MPR, CMR, MPG, CMG, MP, CM, CMX, CMH, FT6, and FT4 have been found to meet the standard criteria for FT1.

Communication cables marked MPP, CMP, MPR, CMR, MPG, CMG, and FT6 have been found to meet the standard criteria for FT4.

Communication cables marked MPP and CMP have been found to meet the standard criteria for FT6.

Figure 14-24. Canadian Electrical Code cable substitution hierarchy per C22.2 #214—Communication Cables.

14.28.5 Plenum Cable

Plenum cable is used in ceilings where the air handling system uses the plenum as the delivery or the return air duct. Because of its flame-resistant and low smoke-emission properties, the special compound used in plenum cable jackets and insulations has been accepted under the provisions of the NEC and classified by Underwriters Laboratories Inc. (UL) for use without conduit in air plenums.

In a typical modern commercial building, cables are installed in the enclosed space between drop ceilings and the floors from which they are suspended. This area is also frequently used as a return air plenum for a building's heating and cooling system. Because these air ducts often run across an entire story unobstructed, they can be an invitation to disaster if fire breaks out. Heat, flames, and smoke can spread rapidly throughout the air duct system and building if the fire is able to feed on combustible materials (such as cable insulations) in the plenum. To eliminate this problem and to keep fumes from entering the air handling system, the NEC requires that conventional cables always be installed in metal conduit when used in plenums.

Plenums, with their draft and openness between different areas, cause fire and smoke to spread, so the 1975 NEC prohibited the use of electrical cables in plenums and ducts unless cables were installed in metal conduit. In 1978, Sections 725-2(b) (signaling cables), 760-4(d) (fire-protection cable), and 800-3(d) (communication/telephone cables) of the NEC allowed that cables "listed as having adequate fire-resistance and low-smoke producing characteristics shall be permitted for ducts, hollow spaces used as ducts, and plenums other than those described in Section 300-22(a)."

While plenum cable costs more than conventional cable, the overall installed cost is dramatically lower because it eliminates the added cost of conduit along with the increased time and labor required to install it.

In 1981 the jacket and insulation compound used in plenum cables was tested and found acceptable under the terms of the NEC and was classified by UL for use without conduit in air return ducts and plenums. Fig. 14-25 shows the UL standard 910 plenum flame test using a modified Steiner tunnel equipped with a special rack to hold test cables.

Virtually any cable can be made in a plenum version. The practical limit is the amount of flammable material in the cable and its ability to pass the Steiner Tunnel Test, shown in Fig. 14-25. Originally plenum cable was all Teflon inside and out. Today most plenum cables have a Teflon core with a special PVC jacket which meets the fire rating. But there are a number of compounds such as Halar® and Solef® that can also be used.

14.28.6 Power Distribution Safety

Electricity kills! No matter how confident we are we must always be careful around electricity. Fibrillation is a nasty and relatively slow death so it is important that Defibrillators are accessible when working around electricity. Table 14-37 displays the small amounts of current that is required to hurt or kill a person.

14.28.6.1 Ground-Fault Interrupters

Ground-fault circuit interrupters (GFCIs) are sometimes called earth leakage or residual-current circuit breakers. GFCIs sense leakage current to earth ground from the hot or neutral leg and interrupt the circuit automatically within 25 ms if the current exceeds 4 to 6 ma. These values are determined to be the maximum safe levels before a human heart goes into ventricular fibrillation. GFCIs do not work when current passes from one line to the other line through a person, for instance. They do not work as a circuit breaker.

One type of GFCI is the core-balance protection device, Fig. 14-26. The hot and neutral power conduc-

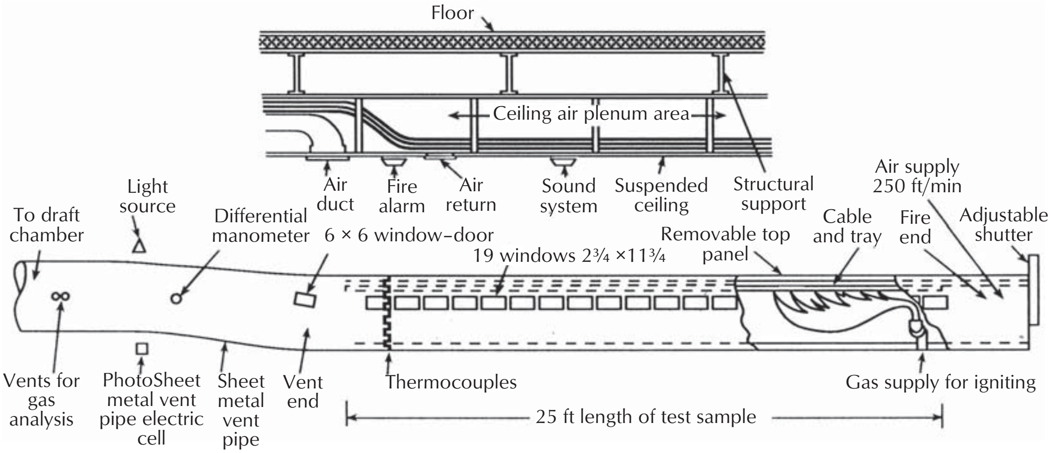

Floor
Ceiling air plenum area
Air duct
Fire alarm
Air return
Sound system
Suspended ceiling
Structural support
Air supply 250 ft/min
Light source
To draft chamber
Differential manometer
6 × 6 window–door
19 windows 2¾ ×11¾
Removable top panel
Cable and tray
Fire end
Adjustable shutter
Vents for gas analysis
PhotoSheet metal vent pipe electric cell
Sheet metal vent pipe
Vent end
Thermocouples
Gas supply for igniting
25 ft length of test sample

Figure 14-25. Plenum cable flame test, UL standard 910.

tors pass through a toroidal (differential) current transformer. When everything is operating properly, the vector sum of the currents is zero. When the currents in the two legs are not equal, the toroidal transformer detects it, amplifies it, and trips an electromagnetic relay. The circuit can also be tested by depressing a test button which unbalances the circuit.

Table 14-37. Physiological Effects of Shock Current on Humans. From Amundson.

Shock Current in mArms	Circuit Resistance at 120 Vac	Physiological Effects
0.5–7 mA	240,000 Ω down to 17,000 Ω	**Threshold of Perception:** Large enough to excite skin nerve endings for a tingling sensation. Average thresholds are 1.1 mA for men and 0.7 mA for women.
1–6 mA	120,000 Ω down to 20,000 Ω	**Reaction Current:** Sometimes called the Surprise current. Usually an involuntary reaction causing the person to pull away from the contact.
6–22 mA	20,000 Ω down to 5400 Ω	**Let-Go Current:** This is the threshold where the person can voluntarily withdraw from the shock current source. Nerves and muscles are vigorously stimulated, eventually resulting in pain and fatigue. Average let-go thresholds are 16 mA for men and 10.5 mA for women. Seek medical attention.
15 mA and above	8000 Ω and below	**Muscular Inhibition:** Respiratory paralysis, pain and fatigue through strong involuntary contractions of muscles and stimulation of nerves. Asphyxiation may occur if current is not interrupted.
60 mA–5 A	2000 Ω down to 24 Ω	**Ventricular Fibrillation:** Shock current large enough to desynchronize the normal electrical activity in the heart muscle. Effective pumping action ceases, even after shock cessation. Defibrillation (single pulse shock) is needed or death occurs.
1 A and above	120 Ω and below	**Myocardial Contraction:** The entire heart muscle contracts. Burns and tissue damage via heating may occur with prolonged exposure. Muscle detachment from bones possible. Heart may automatically restart after shock cessation.

14.28.7 AC Power Cords and Receptacles

Ac power cords, like other cables come with a variety of jacket materials for use in various environments. All equipment should be connected with three-wire cords. Never use ground-lift adapters to remove the ground from any equipment. This can be dangerous, even fatal, if a fault develops inside the equipment, and there is no path to ground.

A common European plug, with a rating of 250 Vac and 10 A, is shown in Fig. 14-27.

The color codes used in North America and Europe for three conductors are shown in Table 14-38.

Cables should be approved to a standards shown in Table 14-39.

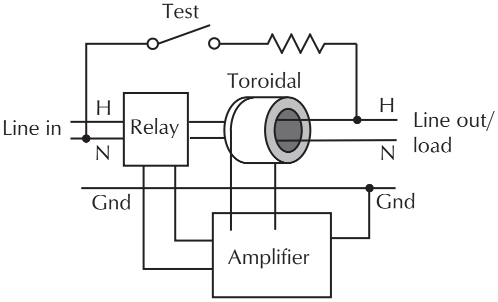


Figure 14-26. Typical ground-fault circuit interrupter.

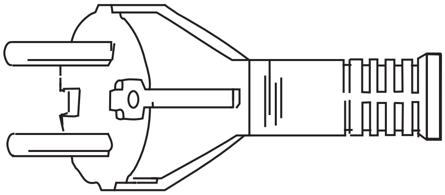


Figure 14-27. Standard international plug.

Table 14-38. Color Codes for Power Supply Cords

Function	North America	CEE and SAA Standard
N—Neutral	White	Light Blue
L—Live	Black	Brown
E—Earth or Ground	Green or Green/Yellow	Green/Yellow

Table 14-39. Approved Electrical Standards

Country	Standard	
United States	UL	Underwriters Laboratory
Canada	cUL	Canadian Underwriters Laboratory
Germany	GS/TUV	German Product Certification Organization
International	IEC	International Electrotechnical Commission

The UL listing signifies that all elements of the cords and assembly methods have been approved by the Underwriters Laboratories, Inc. as meeting their applicable construction and performance standards. UL listed has become a symbol of safety to millions of Americans and their confidence in it results in easier sales of electrical products.

The U.S. NEMA configurations for various voltage and current general purpose plugs and receptacles are shown in Fig. 14-28.

14.28.8 Shielded Power Cable

Shielding the power cord is an effective means of minimizing the error-generating effects of ambient electrical interference. However, the shield effectiveness of most constructions is mostly medium to high-frequency. For instance, braid shields, even high-density high-coverage braids, have little effectiveness below 1000 Hz. So, if the intent is to shield the 50/60 Hz from adjacent cables or equipment, buying a shielded power cord will not be effective. In that case, steel conduit is recommended. But even solid steel conduit, perfectly installed, only has an effectiveness of approximately 30 dB at 50/60 Hz.

The standard power cable shielding consists of aluminum polyester film providing 100% shield coverage and radiation reduction. A spiral-wound drain wire provides termination to ground. These shields are highly effective at high frequencies, generally above

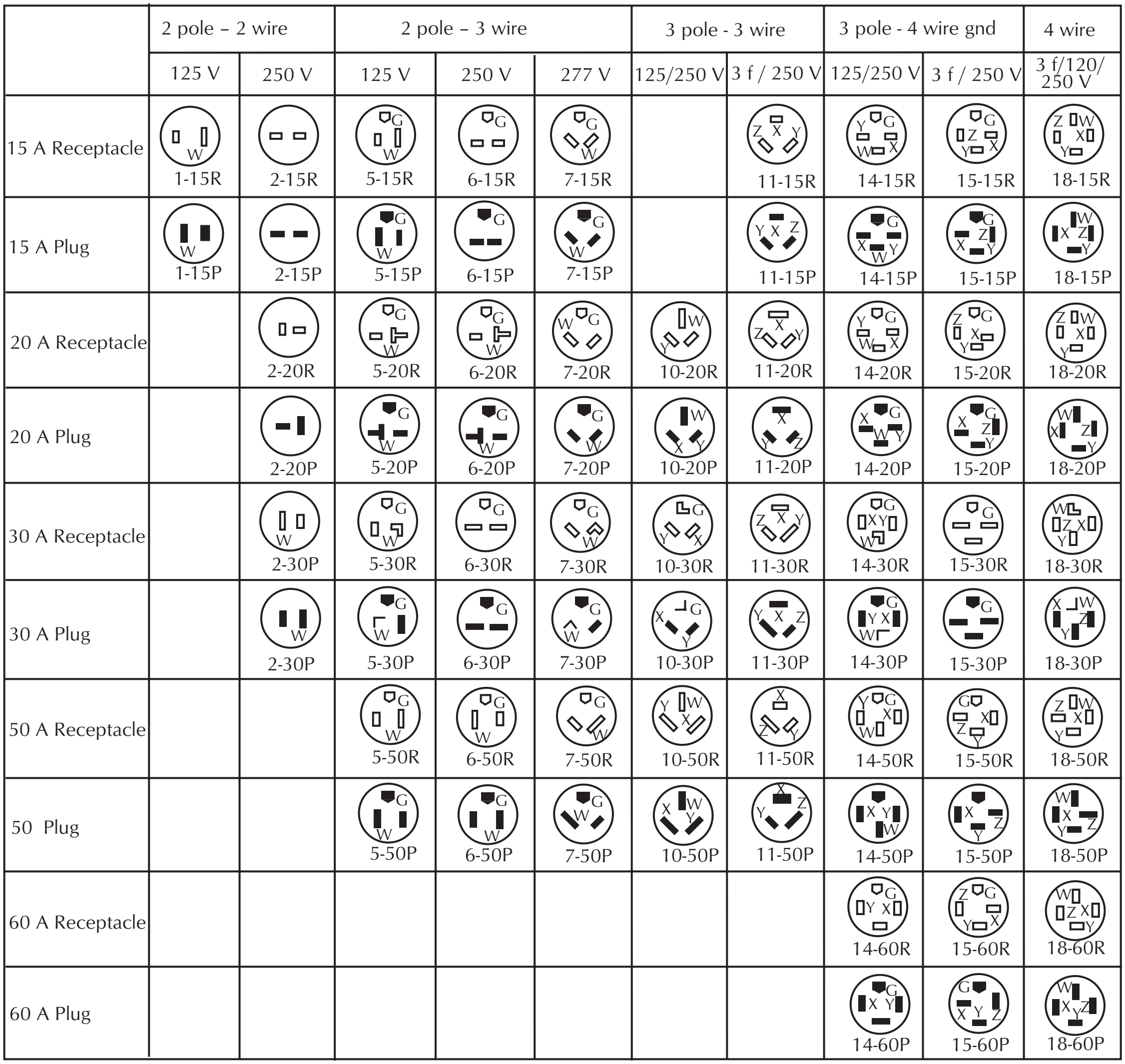

2 pole – 2 wire
2 pole – 3 wire
3 pole - 3 wire
3 pole - 4 wire gnd
4 wire
125 V
250 V
125 V
250 V
277 V
125/250 V
3 f / 250 V
125/250 V
3 f / 250 V
3 f/120/ 250 V
15 A Receptacle
1-15R
2-15R
5-15R
6-15R
7-15R
11-15R
14-15R
15-15R
18-15R
15 A Plug
1-15P
2-15P
5-15P
6-15P
7-15P
11-15P
14-15P
15-15P
18-15P
20 A Receptacle
2-20R
5-20R
6-20R
7-20R
10-20R
11-20R
14-20R
15-20R
18-20R
20 A Plug
2-20P
5-20P
6-20P
7-20P
10-20P
11-20P
14-20P
15-20P
18-20P
30 A Receptacle
2-30P
5-30R
6-30R
7-30R
10-30R
11-30R
14-30R
15-30R
18-30R
30 A Plug
2-30P
5-30P
6-30P
7-30P
10-30P
11-30P
14-30P
15-30P
18-30P
50 A Receptacle
5-50R
6-50R
7-50R
10-50R
11-50R
14-50R
15-50R
18-50R
50 Plug
5-50P
6-50P
7-50P
10-50P
11-50P
14-50P
15-50P
18-50P
60 A Receptacle
14-60R
15-60R
18-60R
60 A Plug
14-60P
15-60P
18-60P
G
W
X
Y
Z

Figure 14-28. NEMA configurations for general-purpose plugs and receptacles.

10 MHz. Power cords used in applications involving extremely high EMI and RFI environments require shield constructions such as Belden Z-Fold™ foil providing 100% coverage, plus another layer of tinned copper braid of 85% coverage, or greater. This provides the maximum shielding available in a flexible power cord.

Shield effectiveness is an important benefit where interference-sensitive electronic devices, such as computer and laboratory test equipment are concerned. However, any designer or installer should realize that the ultimate protection between power cable and other cables or equipment is distance. The inversesquare-law clearly states that double the distance results in four times less interference. Double that distance is sixteen times less, etc.

14.28.9 International Power

Table 14-40 gives the current characteristics for various countries. These are for the majority of the cities in each country, however some countries have different currents in different cities or areas.

Table 14-40. Characteristics of Current in Various Countries

Country	Type of Current	Phases	Voltage	# Wires
Afghanistan	ac 60	1,3	220/380	2,4
Algiers	ac 50	1,3	127/220 220/380	2,4
American Samoa	ac 60	1,3	120/240 240/480	2,3,4
Angola	ac 50	1,3	220/380	2,4
Antiquary	ac 60	1,3	230/400	2,3,4
Argentina	ac 50	1,3	220/380	2,4
Aruba	ac 60	1	115/230	2,3
Australia	ac 50	1,3	240/415	2,3,4
Austria	ac 50	1,3	220/380	3,5
Azores	ac 50	1,3	220/380	2,3,4
Bahamas	ac 60	1,3	120/240	2,3,4
Bahrain	ac 50	1,3	230/400	2,3,4
Balearic Islands	ac 50	1,3	127/220	2,3,4
Bangladesh	ac 50	1,3	220/440	3,4
Barbados	ac 50	1,3	115/230	2,3,4
Belgium	ac 50	1,3	220/380	2,3
Belize	ac 60	1,3	110/220	2,3,4
Benin	ac 50	1,3	220/380	2,4
Bermuda	ac 60	1,3	120/240	2,3,4
Bolivia	ac 50	1,3	230/400	2,4
Botswana	ac 50	1,3	231/400	2,4
Brazil	ac 60	1,3	127/220	2,3,4
Brunei	ac 50	1,3	240/415	2,4
Bulgaria	ac 50	1,3	220/380	2,4
Burkina Faso	ac 50	1,3	220/380	2,4
Burma	ac 50	1,3	230/400	2,4
Burundi	ac 50	1,3	220/380	2,4
Cambodia	ac 50	1,3	120/208	2,4
Cameroon	ac 50	1,3	230/400	2,4
Canada	ac 60	1,3	120/240	3,4
Canary Islands	ac 50	1,3	127/220	2,3,4
Cape Verde	ac 50	1,3	220/380	2,3,4
Cayman Islands	ac 60	1,3	120/240	2,3
Central African Republic	ac 50	1,3	220/380	2,4
Chad	ac 50	1,3	220/380	2,4
Channel Island	ac 50	1,3	240/415	2,4
Chile	ac 50	1,3	220/380	2,3,4
China	ac 50	1,3	220/380	3,4
Columbia	ac 60	1,3	110/220	2,3,4
Comoros	ac 50	1,3	220/380	2,4
Congo	ac 50	1,3	220/380	2,4
Costa Rica	ac 60	1,3	120/240	2,3,4
Cote d'lvoire	ac 50	1,3	220/380	3,4
Cyprus	ac 50	1,3	240/415	2,4
Czechoslovakia	ac 50	1,3	220/380	2,3,4
Denmark	ac 50	1,3	220/380	2,3,4
Djibouti, Rep.	ac 50	1,3	220/380	2,4
Dominica	ac 50	1,3	230/400	2,4
Dominican Republic	ac 60	1,3	110/220	2,3
Ecuador	ac 60	1,3	120/208	2,3,4
Egypt	ac 50	1,3	220/380	2,3,4
El Salvador	ac 60	1,3	115/230	2,3
England	ac 50	1	240/480	2,3
Equatorial Guinea	ac 50	1	220	2
Ethiopia	ac 50	1,3	220/380	2,4
Faeroe Islands	ac 50	1,3	220/380	2,3,4
Fiji	ac 50	1,3	240/415	2,3,4
Finland	ac 50	1,3	220/380	2,4,5
France	ac 50	1,3	220/380 115/230	2,4
French Guiana	ac 50	1,3	220/380	2,3,4
Gabon	ac 50	1,3	220/380	2,4
Gambia, The	ac 50	1,3	220/380	2,4

Table 14-40. Characteristics of Current in Various Countries (Continued)

Country	Type of Current	Phases	Voltage	# Wires
Germany	ac 50	1,3	220/380	2,4
Ghana	ac 50	1,3	220/400	2,3,4
Gibraltar	ac 50	1,3	240/415	2,4
Greece	ac 50	1,3	220/380	2,4
Greenland	ac 50	1,3	220/380	2,3,4
Grenada	ac 50	1,3	230/400	2,4
Guadeloupe	ac 50	1,3	220/380	2,3,4
Guam	ac 60	1,3	110/220	3,4
Guatemala	ac 60	1,3	120/240	2,3,4
Guinea	ac 50	1,3	220/380	2,3,4
Guinea-Bissau	ac 50	1,3	220/380	2,3,4
Guyana	ac 50	1,3	110/220	2,3,4
Haiti	ac 60	1,3	110/220	2,3,4
Honduras	ac 60	1,3	110/220	2,3
Hong Kong	ac 50	1,3	200/346	2,3,4
Hungary	ac 50	1,3	220/380	2,3,4
Iceland	ac 50	1,3	220/380	2,3,4
India	ac 50	1,3	230/400	2,4
Indonesia	ac 50	1,3	127/220	2,4
Iran	ac 50	1,3	220/380	2,3,4
Iraq	ac 50	1,3	220/380	2,4
Ireland	ac 50	1,3	220/380	2,4
Isle of Man	ac 50	1,3	240/415	2,4
Israel	ac 50	1,3	230/400	2,4
Italy	ac 50	1,3	127/220	2,4
Jamaica	ac 50	1,3	110/220	2,3,4
Japan	ac 60	1,3	100/200	2,3
Jerusalem	ac 50	1,3	220/380	2,3,4
Jordan	ac 50	1,3	220/380	2,3,4
Kenya	ac 50	1,3	240/415	2,4
Korea	ac 60	1	110	2
Kuwait	ac 50	1,3	240/415	2,4
Laos	ac 50	1,3	220/380	2,4
Lebanon	ac 50	1,3	110/190	2,4
Lesotho	ac 50	1,3	220/380	2,4
Liberia	ac 60	1,3	120/240	2,3,4
Libya	ac 50	1,3	127/220	2,4
Luxembourg	ac 50	1,3	220/380	3,4,5
Macao	ac 50	1,3	200/346	2,3
Madagascar	ac 50	1,3	220/380	2,4
Madeira	ac 50	1,3	220/380	2,3,4
Malawi	ac 50	1,3	230/400	3,4
Malaysia	ac 50	1,3	240/415	2,4

Table 14-40. Characteristics of Current in Various Countries (Continued)

Country	Type of Current	Phases	Voltage	# Wires
Maldives	ac 50	1,3	230/400	2,4
Mali, Rep. of	ac 50	1,3	220/380	2,4
Malta	ac 50	1,3	240/415	2,4
Martinique	ac 50	1,3	220/380	2,3,4
Mauritania	ac 50	1,3	220	2,3
Mauritius	ac 50	1,3	230/400	2,4
Mexico	ac 60	1,3	127/220	2,3,4
Monaco	ac 50	1,3	127/220	2,4
Montserrat	ac 60	1,3	230/400	2,4
Morocco	ac 50	1,3	127/220	2,4
Mozambique	ac 50	1,3	220/380	2,3,4
Namibia	ac 50	1,3	220/380	2,4
Nassau	ac 60	1,3	120/240	2,3,4
Nepal	ac 50	1,3	220/440	2,4
Netherlands	ac 50	1,3	220/380	2,3
Netherlands, Antilles	ac 50	1,3	127/220	2,3,4
New Caledonia	ac 50	1,3	220/380	2,3,4
New Zealand	ac 50	1,3	230/400	2,3,4
Nicaragua	ac 60	1,3	120/240	2,3,4
Niger	ac 50	1,3	220/380	2,3,4
Nigeria	ac 50	1,3	230/415	2,4
Northern Ireland	ac 50	1,3	220/380 240/480	2,4
Norway	ac 50	1,3	230	2,3
Okinawa	ac 60	1	120/240	2,3
Oman	ac 50	1,3	240/415	2,4
Pakistan	ac 50	1,3	220/380	2,3,4
Panama	ac 60	1,3	110/220	2,3
Papua New Guinea	ac 50	1,3	240/415	2,4
Paraguay	ac 50	1,3	220/380	2,4
Peru	ac 50	1,3	220	2,3
Philippines	ac 60	1,3	110/220	2,3
Poland	ac 50	1,3	220/380	2,4
Portugal	ac 50	1,3	220/380	2,3,4
Puerto Rico	ac 60	1,3	120/240	2,3,4
Qatar	ac50	1,3	240/415	2,3,4
Romania	ac 50	1,3	220/380	2,4
Russia	ac 50	1,3	220/380	
Rwanda	ac 50	1,3	220/380	2,4
Saudi Arabia	ac 60	1,3	127/220	2,4
Scotland	ac 50	1,3	240/415	2,4
Senegal	ac 50	1,3	127/220	2,3,4
Seychelles	ac 50	1,3	240	2,3

Table 14-40. Characteristics of Current in Various Countries (Continued)

Country	Type of Current	Phases	Voltage	# Wires
Sierra Leone	ac 50	1,3	230/400	2,4
Singapore	ac 50	1,3	230/400	2,4
Somalia	ac 50	1,3	230	2,3
South Africa	ac 50	1,3	220/380	2,3,4
Spain	ac 50	1,3	127/220 220/380	2,3,4
Sri Lanka	ac 50	1,3	230/400	2,4
St. Kitts and Nevis	ac 60	1,3	230/400	2,4
St. Lucia	ac 50	1,3	240/416	2,4
St. Vincent	ac 50	1,3	230/400	2,4
Sudan	ac 50	1,3	240/415	2,4
Suriname	ac 60	1,3	127/220	2,3,4
Swaziland	ac 50	1,3	230/400	2,4
Sweden	ac 50	1,3	220/380	2,3,4,5
Switzerland	ac 50	1,3	220/380	2,3,4
Syria	ac 50	1,3	220/380	2,3
Tahiti	ac 60	1,3	127/220	2,3,4
Taiwan	ac 60	1,3	110/220	2,3,4
Tanzania	ac 50	1,3	230/400	2,3,4
Thailand	ac 50	1,3	220/380	2,3,4
Togo	ac 50	1,3	220/380	2,4
Tonga	ac 50	1,3	240/415	2,3,4
Trinidad and Tobago	ac 60	1,3	115/230 230/400	2,3,4
Tunisia	ac 50	1,3	127/220 220/380	2,4
Turkey	ac 50	1,3	220/380	2,3,4
Uganda	ac 50	1,3	240/415	2,4
United Arab Emirates	ac 50	1,3	220/415	2,3,4
United Kingdom	ac 50	1	240/480	2,3
United States	ac 60	1,3	120/240	3,4
Uruguay	ac 50	1,3	220	2,3
Venezuela	ac 60	1,3	120/240	2,3,4
Vietnam	ac 50	1,3	220/380	2,4
Virgin Islands (American)	ac 60	1,3	120/240	2,3,4
Wales	ac 50	1,3	240/415	2,4
Western Samoa	ac 50	1,3	230/400	2,4
Yemen Arab Republic	ac 50	1,3	230/400	2,4
Yugoslavia	ac 50	1,3	220/380	2,4
Zaire, Rep. of	ac 50	1,3	220/380	2,3,4
Zambia	ac 50	1,3	220/380	2,4
Zimbabwe	ac 50	1,3	220/380	2,3,4

Acknowledgments

I wish to acknowledge Martin Van Der Burgt (Product Development Manager) and Carl Dole and Jason R. Kovaric (Applications Engineers), both with Belden Americas Division.

Steve Lampen, Belden

Bibliography

"Conduit Capacity Chart," Belden Technical Papers.

"Characteristic Impedance of Cables at High and Low Frequencies," Belden Technical Papers.

Marty Van Der Burgt, "Solving Signal Problems," Belden Technical Papers.

Paul Vanderlaan, "Installation Effects Upon Alien Crosstalk and Equal Level Far End Crosstalk," Belden Technical Paper, October 1999.

Paul Vanderlaan, "Impedance/Return Loss," Belden Technical Papers.

R. A. Greiner, "Cables and the Amp," *Audio*, August 1989. Parts copied with permission.

Steve Lampen, *The Audio-Video Cable Installer's Pocket Guide*, New York:McGraw-Hill.

Steve Lampen, "Cable and the Cutting Edge," Belden Technical Papers, July 1997.

Steve Lampen, "How Wire Fails," Belden Technical Papers, July 1997.

Steve Lampen, "Video and UTP," Belden Technical Papers, May 2000.

Steve Lampen, "What Is RG-6 Anyway," CE Pro, July 2000.

Steve Lampen, *Wire, Cable, and Fiber Optics for Video and Audio Engineers*, New York: McGraw Hill.

Steve Lampen, "Wire, Cable, and Fiber Optics," Belden Technical Papers, June 2000.

"Characteristic Impedance of Cables at High and Low Frequencies," Engineering Departmen,t Belden Electronics Division.

"Choosing the Correct Coaxial Cable for CCTV Applications," Technical Bulletin #*TB3-001*, West Penn Wire

"Data Grade Fire Alarm Cables," West Penn Technical Wire.

"Differences in Coaxial Cable Construction," West Penn Technical Wire.

"Electric Current Abroad," 1991 Edition, U.S. Department of Commerce.

Chapter 15

Transmission Techniques: Fiber Optics

by Ron Ajemian

15.1 History 451
15.2 Advantages of Using Fiber Optics for Audio 452
  15.2.1 Applications for Audio 452
15.3 Physics of Light 452
15.4 Fiber Optics 456
  15.4.1 Types of Fiber 456
  15.4.2 Characteristics of Typical Fibers. 457
  15.4.3 Signal Loss 458
    15.4.3.1 Fiber Optic Transmission Loss (FOTL) 458
    15.4.3.2 Optical Fiber Loss 458
    15.4.3.3 Attenuation Measurement 461
    15.4.3.4 Advancements in OTDR Testing 462
15.5 Sources 463
  15.5.1 LEDs 463
  15.5.2 Laser Diodes 465
  15.5.3 Superluminescent Diodes (SLDs) 466
  15.5.4 Vertical Cavity Surface Emitting Laser (VCSEL) 466
  15.5.5 LED and Laser Characteristics 466
    15.5.5.1 Output Power 466
    15.5.5.2 Output Pattern 466
    15.5.5.3 Wavelength 466
    15.5.5.4 Speed 467
    15.5.5.5 Lifetime 467
    15.5.5.6 Safety 467
15.6 Detectors 467
  15.6.1 Quantum Efficiency (η) 468
  15.6.2 Noise 468
  15.6.3 Bandwidth 468
15.7 Transmitter/Receiver Modules 469
15.8 Transceivers and Repeaters 469
  15.8.1 Demand on Gigabit Optical Transceivers 469
15.9 The Fiber Optic Link 470
  15.9.1 Fiber Optic Link Design Considerations. 470
  15.9.2 Passive Optical Interconnections. 470
  15.9.3 Fiber Optic Connectors 471
    15.9.3.1 Biconic Connector 471

- 15.9.3.2 SMA Connector . . . 471
- 15.9.3.3 FC/PC Connector . . . 471
- 15.9.3.4 ST Connector . . . 472
- 15.9.3.5 SC Connector . . . 472
- 15.9.3.6 D4 Connector . . . 472
- 15.9.3.7 FDDI Connector . . . 472
- 15.9.3.8 SFF (Small Form Factor) Connectors . . . 472
- 15.9.3.9 LC Connector . . . 472
- 15.9.3.10 MT-RJ Connector . . . 473
- 15.9.3.11 OpticalCon® . . . 473
- 15.9.3.12 Toslink . . . 473
- 15.9.3.13 Military Grade Connectors . . . 473
- 15.9.4 Connector Installation . . . 473
  - 15.9.4.1 Current Fiber Optic Connectors for Glass Optical Fibers (GOF) . . . 473
  - 15.9.4.2 LC Versus Other Types of Fiber Optic Connectors (ST, SC, etc.) . . . 474
  - 15.9.4.3 Fiber Optic Connector Termination Advancements . . . 474
  - 15.9.4.4 Fiber Optic Splices . . . 474
- 15.10 Couplers . . . 475
- 15.11 Fiber Optic System Design . . . 475
  - 15.11.1 System Specifications . . . 475
    - 15.11.1.1 Transmitter Specifications . . . 476
    - 15.11.1.2 Light Source Specifications . . . 476
    - 15.11.1.3 Fiber Specifications . . . 476
    - 15.11.1.4 Cable Specifications . . . 476
    - 15.11.1.5 Detector Specifications . . . 476
    - 15.11.1.6 Receiver Specifications . . . 476
  - 15.11.2 Design Considerations . . . 476
  - 15.11.3 Design Procedures . . . 477
    - 15.11.3.1 Fiber Optic System Attenuation . . . 477
    - 15.11.3.2 Additional Losses . . . 477
- 15.12 Fiber Optic Considerations . . . 479
- 15.13 Glossary of Fiber Optic Terms . . . 479
- Bibliography . . . 484

15.1 History

Fiber optics is the branch of optical technology concerned with the transmission of light through fibers made of transparent materials, such as glass, fused silica, or plastic, to carry information.

Fiber optics has been used by the telephone industry for over thirty years, and has proved itself as being the transmission medium for communications. Past history shows audio follows the telephone industry, therefore fiber optics will soon be a force in audio.

The founder of fiber optics was probably the British physicist, John Tyndall. In 1870 Tyndall performed an experiment before the Royal Society that showed light could be bent around a corner as it traveled in a rush of pouring water. Tyndall aimed a beam of light through the spout along with the water and his audience saw the light follow a zigzag path inside the curved path of the water. His experiment utilized the principle of total internal reflection, which is also applied in today's optical fibers.

About ten years later, William Wheeler, an engineer from Concord, Massachusetts invented a scheme for piping light through buildings. He used a set of pipes with a reflective lining and diffusing optics to transmit light (bright electric arc) through a building, then diffuse it into other rooms. Although Wheeler's light pipes probably didn't reflect enough light to illuminate the rooms, his idea kept coming up again and again until it finally coalesced into the optical fiber.

At about the same time Alexander Graham Bell invented the photophone, Fig.15-1. Bell demonstrated that a light ray could carry speech through the air. This was accomplished by a series of mirrors and lenses directing light onto a flat mirror attached to a mouthpiece. Speech vibrating the mirror caused the light to modulate. The receiver included a selenium diode detector whose resistance varied with the intensity of light striking it. Thus the modulated light (sunlight, etc.) striking the selenium detector varied the amount of current through the receiver and reproduced speech that could be transmitted over distances of approximately 200 meters.

In 1934, an American, Norman R. French, while working with AT&T, received a patent for his optical telephone system. French's patent described how speech signals could be transmitted via an optical cable network. Cables were to be made out of solid glass rods or a similar material with a low attenuation coefficient at the operating wavelength.

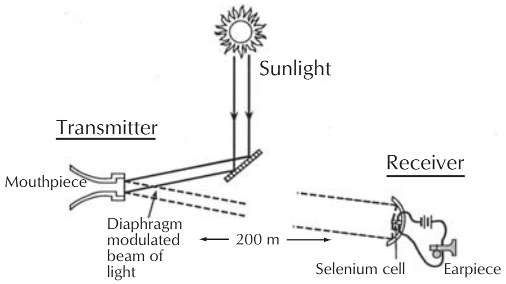


Figure 15-1. Alexander Graham Bell's Photophone.

Interest in glass waveguides increased in the 1950s, when research turned to glass rods for unmodulated transmission of images. One result was the invention of the fiber scope, widely used in the medical field for viewing the internal parts of the body. In 1956 Brian O'Brien, Sr., in the United States, and Harry Hopkins and Narinder Kapany, in England, found the way to guide light. The key concept was making a two-layer fiber. One layer was called the core the other layer was called the cladding (see section on light). Kapany then coined the term *fiber optics.*

An efficient light source was needed but it wasn't until 1960 when the first laser light was invented that it became available. A Nobel Prize was awarded to Arthur Schawlow and Charles H. Townes of Bell Laboratories for developing the laser, which was first successfully operated by Theodor H. Maiman of Hughes Research Laboratory. The manufacturing process of lasers from semiconductor material was recognized in 1962. At the same time semiconductor photodiodes were developed for receiver elements. Now the only thing left was to find a suitable transmission medium.

Then in 1966 Charles H. Kao and George A. Hockham, of Standard Telecommunication Labs, England, published a paper proposing that optical fibers could be used as a transmission medium if their losses could be reduced to 20 dB/km. They knew that high losses of over 1000 dB/km were the result of impurities in the glass, not of the glass itself. By reducing these impurities a low-loss fiber could be produced for telecommunications.

Finally in 1970, Robert Maurer and associates at Corning Glass Works, New York, developed the first fiber with losses under 20 dB/km, and by 1972 lab samples were revealed as low as 4 dB/km. Since then the Corning Glass Works and Bell Telephone Labs of the United States; Nippon Sheet Glass Company and Nippon Electric Company of Japan; and at AEG-Telefunken, Siemens and Halske in Germany, have developed glass fibers with losses at about 0.2 dB/km. There is also some plastic materials as well as glass being used for shorter distances.

The practical use of fiber optics for communications began in the mid- and late 1970s with test trials. However, the popularization of fiber optics wasn't until the 1980 Winter Olympics at Lake Placid, New York, when the joint effort of New York Telephone, AT&T, Western Electric, and Bell Labs installed a fiber optic system. Its purpose was to transform the Lake Placid telephone facility into a professional communications center capable of handling a wide range of telecommunications services necessary to support the Olympic events. Today fiber optics is an established technology.

15.2 Advantages of Using Fiber Optics for Audio

There are at least four advantages in using fiber over hardwired systems. One is the superb performance in transmission, allowing extremely large bandwidths and low loss which minimizes the need for preamplifying a signal for long haul applications. Digital data can be easily transmitted with rates of 100 Mb/s or higher showing more information handling capability and greater efficiency. Since the optical fiber is nonmetallic (made of glass, plastic, etc.), it is immune to problems caused by electromagnetic interference (EMI) and radio frequency interference (RFI). Also the problem of crosstalk is eliminated—a quality advantage.

With optical fiber one no longer needs to worry about impedance matching, electrical grounding or shorting problems, or no ground loops. Safety is an important feature of fiber optics because a broken cable will not spark, possibly causing shock or an explosion in a dangerous environment.

Another plus is fiber optic cable weighs about 9 lbs/1000 ft and takes up less space than wire, useful especially when running in conduits. Cost is now less than or comparable to copper. And finally an optical fiber system cannot be easily tapped, which allows for better security.

15.2.1 Applications for Audio

Telephone companies have many fiber links which can connect Japan and Europe to the United States. Think of the many possibilities of doing a multitrack recording from many different places all over the world over a fiber optic cable without worrying about SNR, interference, distortion, etc. Top-of-the-line compact disc and DAT players already provide an optical fiber link output. Also, there are companies like Klotz Digital of Germany and Wadia Digital Corporation of the United States who are manufacturing fiber optic digital audio links, which employ an AES/EBU input and output at each end.

Many recording studios are located in high rise apartment buildings. A perfect application of a digital audio fiber optic link is to connect, for instance, studio A which is located on the 21st floor, to studio B which is located on the 24th floor. This is ideal because the user doesn't have to worry about noise and interference caused by fluorescent lighting and elevator motors, to name a few. Another perfect use is to connect MIDI stations together.

Another recent advance is a recording studio can record in real time by using DWDM (dense wavelength division multiplexing) lasers and erbium doped optical fibers to send the AES3 audio channels over the Atlantic or Pacific Ocean and then to the appropriate recording studio. The Internet is also being used to establish a fiber optic end-to-end recording session.

15.3 Physics of Light

Before discussing optical fiber, we must understand the physics on light.

Light. Light is electromagnetic energy, as are radio waves, x-rays, television, radar, and electronic digital pulses. The frequencies of light used in fiber optic data transmission are around 200 THz–400 THz (400×10^{12}), several orders of magnitude higher on the electromagnetic energy spectrum than the highest radio waves, see Fig. 15-2. Wavelength, a more common way of describing light waves, are correspondingly shorter than radio wavelengths. Visible light, with wavelengths from about 400 nm for deep violet to 750 nm for deep red, is only a small portion of the light spectrum. While fiber optic data transmission sometimes uses visible light in the 600 nm to 700 nm range, the near infrared region extending from 750 nm to 1550 nm is of greater interest because fibers propagate the light of these wavelengths more efficiently.

The main distinction between different waves lies in their frequency or wavelength. Frequency, of course, defines the number of sine-wave cycles per second and is expressed in hertz (Hz). Wavelength is the distance between the same points on two consecutive waves (or it is the distance a wave travels in a single cycle). Wavelength and frequency are related. The wavelength (λ) equals

$$\lambda = \frac{v}{f} \tag{15-1}$$

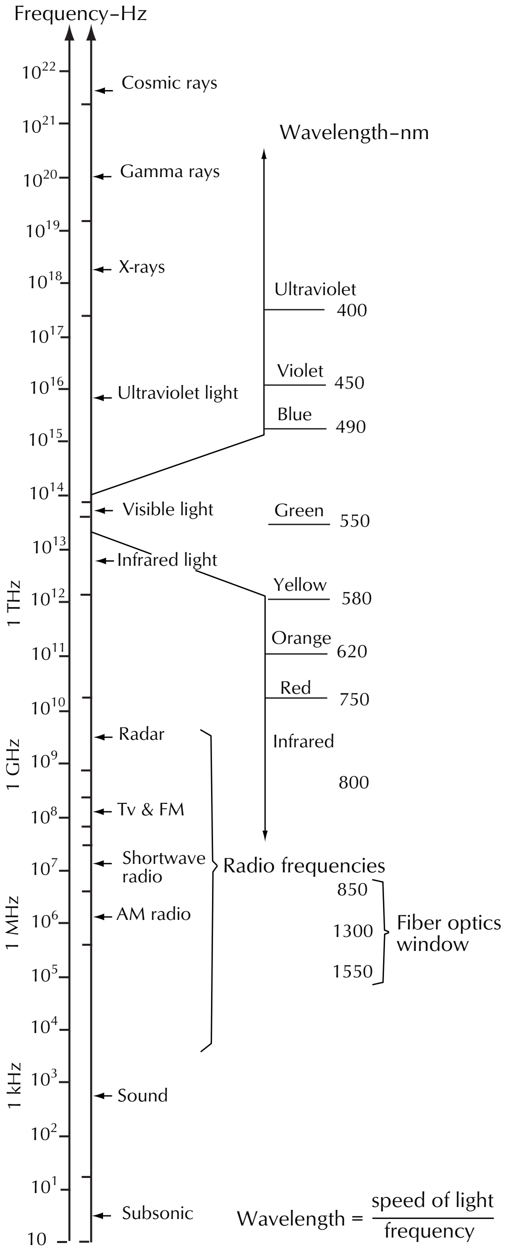

Frequency–Hz
10^{22}
Cosmic rays
10^{21}
Wavelength–nm
10^{20}
Gamma rays
10^{19}
X-rays
10^{18}
Ultraviolet
400
10^{17}
Violet
450
10^{16}
Ultraviolet light
Blue
490
10^{15}
10^{14}
Visible light
Green
550
10^{13}
Infrared light
Yellow
580
1 THz
10^{12}
Orange
620
10^{11}
Red
750
10^{10}
Radar
Infrared
1 GHz
10^{9}
800
10^{8}
Tv & FM
10^{7}
Shortwave radio
Radio frequencies
850
1300
1550
Fiber optics window
1 MHz
10^{6}
AM radio
10^{5}
10^{4}
1 kHz
10^{3}
Sound
10^{2}
10^{1}
Subsonic
10
Wavelength = speed of light / frequency

Figure 15-2. The electromagnetic spectrum.

The velocity of electromagnetic energy in free space is generally called *the speed of light, (*186,000 mi/s [300,000 km/s]). The equation clearly shows that the higher the frequency, the shorter the wavelength.

Light travels slower in other media than a vacuum, and different wavelengths travel at different speeds in the same medium. When light passes from one medium to another, it changes speed, causing a deflection of light called refraction. A prism demonstrates this principle. White light entering a prism is composed of all colors which the prism refracts. Because each wavelength changes speed differently, each is refracted differently, therefore the light emerges from the prism divided into the colors of the visible spectrum, as shown in Fig.15-3.

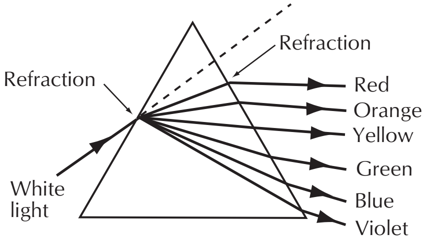

Refraction
Refraction
Red
Orange
Yellow
Green
Blue
Violet
White light

Figure 15-3. Light prism.

The Particle of Light. Light and electrons both exhibit wave- and particlelike traits. Albert Einstein theorized that light could interact with electrons so that the light itself might be considered as bundles of energy or quanta (singular, quantum). This helped explain the photoelectric effect.

In this concept, light rays are considered to be particles that have a zero rest mass called photons.

The energy contained in a photon depends on the frequency of the light and is expressed in Planck's Law, as

$$E = hf \tag{15-2}$$

where,

E is the energy in watts,

h is Planck's constant, equal to 6.624×10^{-34} joule-second,

f is its frequency.

As can be seen from this equation, light energy is directly related to frequency (or wavelength). As the frequency increases, so does the energy, and vice versa. Photon energy is proportional to frequency. Because most of the interest in photon energy is in the part of the spectrum measured in wavelength, a more useful equation which gives energy in electron volts when wavelength is measured in micrometers (μm) is

$$E \text{ in ev} = \frac{1.2406}{\lambda \text{ in } \mu\text{m}}. \tag{15-3}$$

Treating light as both a wave and a particle aids investigation of fiber optics. We switch back and forth between the two descriptions, depending on our needs. For example the characteristics of many optical fibers vary with wavelength, so the wave description is used. On the other hand, the emission of a light by a source, a light emitting diode (LED), or its absorption by a positive-intrinsic-negative detector (PIN), is best treated by particle theory.

Light Rays. The easiest way to view light in fiber optics is by using light ray theory, where the light is treated as a simple ray drawn by a line. The direction of propagation is shown on the line by an arrow. The movement of light through the fiber optic system can be analyzed with simple geometry. This approach simplifies the analysis and makes the operation of an optical fiber simple to understand.

Refraction and Reflection. The index of refraction (n) is a dimensionless number expressing the ratio of the velocity of light in free space (c) to its velocity in a specific medium (v)

$$n = \frac{c}{v} \tag{15-4}$$

The following are typical indices of refraction:

Vacuum	1.0
Air	1.0003 (generalized to 1)
Water	1.33
Fused Quartz	1.46
Glass	1.5
Diamond	2.0
Gallium Arsenide	3.35
Silicon	3.5
Aluminum Gallium Arsenide	3.6
Germanium	4.0

Although the index of refraction is affected by light wavelength, the influence of wavelength is small enough to be ignored in determining the refractive indices of optical fibers.

Refraction of a ray of light as it passes from one material to another depends on the refractive index of each material. In discussing refraction, three terms are important. The *normal* is an imaginary line perpendicular to the interface of the two materials. The *angle of incidence* is the angle between the incident ray and the normal. The *angle of refraction* is the angle between the normal and the refracted ray.

When light passes from one medium to another that has a higher refractive index, the light is refracted toward the normal as shown in Fig. 15-4A. When the index of the first material is higher than that of the second, most of the light is refracted away from the normal, Fig. 15-4B. A small portion is reflected back into the first material by Fresnel reflection. The greater the difference in the indices of two materials the greater the reflection. The magnitude of the Fresnel reflection at the boundary between any two materials is approximately

$$R = \left(\frac{n_1 - n_2}{n_1 + n_2}\right)^2 \tag{15-5}$$

where,

R is the Fresnel reflection,

n_1 is the index of refraction of material 1,

n_2 is the index of refraction of material 2.

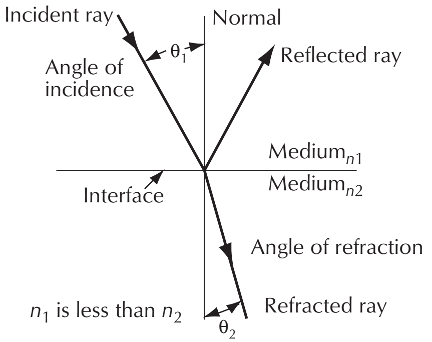
In decibels, this loss of transmitted light is

$$L_F = -10\log(1 - R)\text{ dB}. \tag{15-6}$$

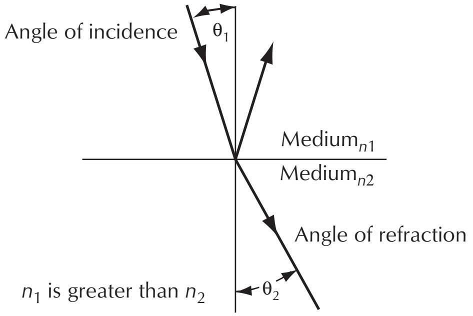
As the angle of incidence increases, the angle of refraction approaches 90° with the normal. The angle of incidence that yields a 90° angle of refraction is called the critical angle, Fig. 15-4C. If the angle of incidence is increased past the critical, the light is totally reflected back into the first material and does not enter the second material and the angle of reflection equals the angle of incidence, Fig. 15-4D.

A single optical fiber is comprised of two concentric layers. The inner layer, the *core,* contains a very pure glass (very clear glass); it has a refractive index higher than the outer layer, or *cladding,* which is made of less pure glass (not so clear glass). Fig. 15-5 shows the arrangement. As a result, light injected into the core and striking the core-to-cladding interface at an angle greater than the critical is reflected back into the core. Since the angles of incidence and reflection are equal, the ray continues zigzagging down the length of the core by total internal reflection, as shown in Fig. 15-6. The light is trapped in the core, however, the light striking the interface at less than the critical angle passes into the cladding and is lost. The cladding is usually surrounded by a third layer, the buffer, whose purpose is to protect the optical properties of the cladding and core.

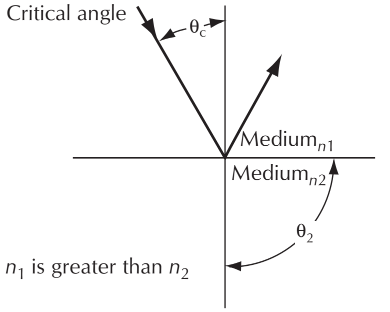
Total internal reflection forms the basis for light propagation in optical fiber. Most analyses of light propagation in a fiber evaluate meridional rays—those which pass through the fiber axis each time they are reflected. To help you to understand how an optical fiber works, let us look at Snell's Law which describes


Incident ray
Normal
θ_1
Angle of incidence
Reflected ray
Medium$_{n1}$
Medium$_{n2}$
Interface
Angle of refraction
n_1 is less than n_2
Refracted ray
θ_2

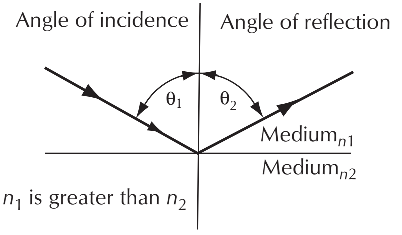
A. Light is bent toward the normal.


Angle of incidence
θ_1
Medium$_{n1}$
Medium$_{n2}$
Angle of refraction
n_1 is greater than n_2
θ_2

B. Light is bent away from the normal.


Critical angle
θ_c
Medium$_{n1}$
Medium$_{n2}$
θ_2
n_1 is greater than n_2

C. Light does not enter the second material.


Angle of incidence
Angle of reflection
θ_1
θ_2
Medium$_{n1}$
Medium$_{n2}$
n_1 is greater than n_2

D. When the angle of incidence is more than the critical angle, light is reflected.

n_1, n_2 are the refractive indices.

Figure 15-4. Refraction and reflection.

Example: A 50/125 fiber nomenclature indicates both the outside diameter of the core (50 microns) and the cladding (125 microns)

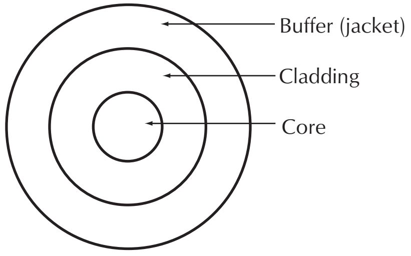

Buffer (jacket)
Cladding
Core

Figure 15-5. Optical fiber cross section.

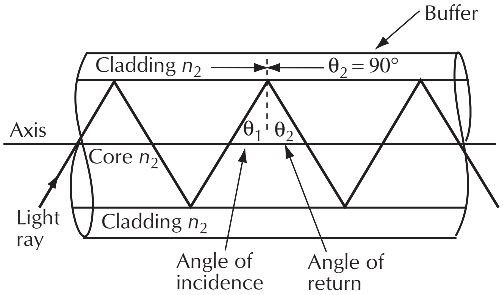

Buffer
Cladding n_2
$\theta_2 = 90°$
Axis
θ_1
θ_2
Core n_2
Light ray
Cladding n_2
Angle of incidence
Angle of return

Figure 15-6. Light guided through an optical fiber.

the relationship between incident and reflected light as shown in Fig. 15-6.

Snell's Law equation is

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \qquad (15\text{-}7)$$

where,

n_1 is the refractive index of the core,
n_2 is the refractive index of the cladding,
θ_1 is the angle of incidence,
θ_2 is the angle of reflection.

The critical angle of incidence, θ_c, (where $\theta_2 = 90°$) is

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \qquad (15\text{-}8)$$

At angles greater than θ_c, the light is reflected. Because reflected light means that n_1 and n_2 are equal (since they are in the same material), θ_1 and θ_2, the angles of incidence and reflection are equal. These simple principles of refraction and reflection form the basis of light propagation through an optical fiber.

Fibers also support skew rays, which travel down the core without passing through the fiber axis. In a straight fiber, the patch of a skew ray is typically helical. Because skew rays are very complex to analyze, they are usually not included in practical fiber analysis. The exact characteristics of light propagation depend on the fiber size, construction, and composition, and on the nature of the light source injected.

Fiber performance and light propagation can be reasonably approximated by considering light as rays. However, more exact analysis must deal in field theory and solutions to Maxwell's electromagnetic equations. Maxwell's equations show that light does not travel randomly through a fiber; it is channeled into *modes*, which represent allowed solutions to electromagnetic field equations. In simple terms, a mode is a possible path for a light traveling down a fiber.

The characteristics of the glass fiber, in an extreme sense, can be compared to light as seen through crystal clear water, turbid water, and water containing foreign objects. These conditions are characteristics of water and have quite different effects on light traveling (propagating) through them. The glass fibers are no different, splices, breaks, boundary distortion, bubbles, core out-of-round, etc., all influence the amount of light that reaches the distant end. The main objective is to receive maximum intensity with little or no distortion.

15.4 Fiber Optics

15.4.1 Types of Fiber

Optical fibers are usually classified by their refractive index profiles and their core size. There are three main types of fibers:

1. Single mode.
2. Multimode stepped index.
3. Multimode graded index.

Single Mode Fiber. Single mode fiber contains a core diameter of 8 to 10 microns, depending on the manufacturer. A highly concentrated source such as a laser or high-efficient LED must be used to produce a single mode for radiation into the fiber. The index of refraction in single mode fiber is very low because the highly concentrated beam and extremely small core prevent blossoming (officially referred to as scattering) of the ray.

The small core tends to prevent the entry of extraneous modes into the fiber, as illustrated in Fig. 15-7. Loss in a single mode fiber is very low and permits the economy of longer repeater (telephone amplifier) spacing. This optical fiber has the capability of propagating 1310 nm and 1550 nm wavelengths. It is well suited for intracity and intercity applications where long repeater spacing is desired.

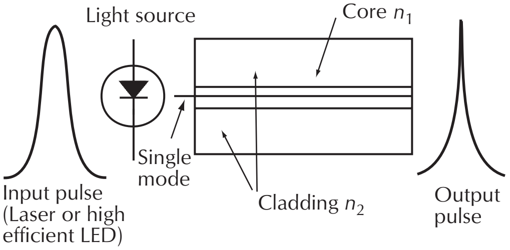


Figure 15-7. Single mode fiber.

Multimode Step Index Fiber. The production of optical fiber includes layer deposition of core glass inside a started tube. If the glass core layers exhibit the same optical properties the fiber is classed a *step index fiber*. The core layers contain uniform transmission characteristics. The fanout of the rays and their refraction at the core-clad boundary give them the appearance of stepping through the glass, Fig. 15-8. Notice also that as the individual rays step their way through, some travel farther and take longer to reach the far end; the reason for the rounded output pulse shown. This optical fiber requires repeaters-regenerators located at short intervals.

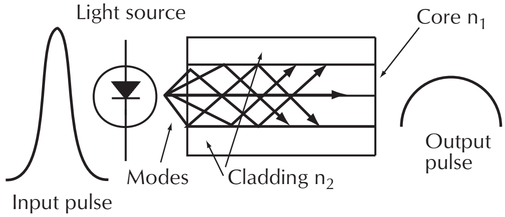


Figure 15-8. Multimode step index fiber.

All rays, or modes, arriving in unison will produce the most exact and strongest replica of the input; this is the objective. For an optical fiber to be most useful in communications, the modes must be channeled through the core in a controlled manner so they all arrive at nearly the same instant.

Multimode Graded Index Fiber. The process of manufacturing graded index fiber involves depositing different grades of glass in the starting tube to provide a core with various transmission characteristics; the outer portion does not impede the passage of modes as much as the center.

In graded index fiber, the core axis contains a higher-density glass of slow wave (ray, mode) propagation in this path for coordination with arrival of the waves in the longest path. The grades of core glass deposited from axis to perimeter are progressively less impeding to let all waves arrive in unison and greatly increase the received intensity (power).

Notice in Fig. 15-9 how each mode is bent (and slowed) in proportion to its entry point in the optical fiber, keeping them in phase. When the rays arrive in phase their powers add. This technique provides maximum signal strength over the greatest distance without regeneration because out-of-phase modes subtract from the total power.

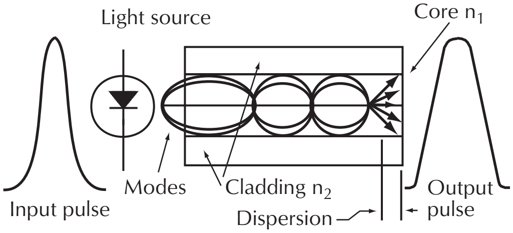


Figure 15-9. Multimode graded index fiber.

15.4.2 Characteristics of Typical Fibers

Table 15-1 gives the characteristics of typical fiber optic cable.

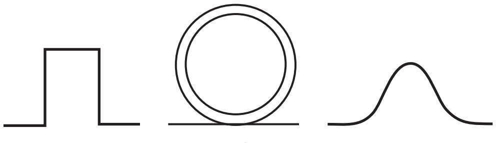
Table 15-1. Characteristics of Typical Cables

Type	Core Dia. (μm)	Cladding Dia. (μm)	Buffer Dia. (μm)	*NA*	Bandwidth MHz-km	Attenuation dB/km
Single mode	8	125	250		6 ps/km*	0.5
at 1300 nm	5	125	250		4 ps/km*	0.4
Graded index	50	125	250	0.20	400	3
at 850nm	62.5	125	250	0.275	150	3
	85	125	250	0.26	200	3
	100	140	250	0.30	150	4
Step index	200	380	600	0.27	25	6
at 850 nm	300	440	650	0.27	20	6
PCS†	200	350	—	0.30	20	10
at 790 nm	400	450	—	0.30	15	10
	600	900	—	0.40	20	6
Plastic	—	750	—	0.50	20	400
at 650 nm	—	1000	—	0.50	20	400

*Dispersion per nanometer of source width.
†PCS (Plastic-clad silica: plastic cladding and glass core).
(Courtesy AMP Incorporated)

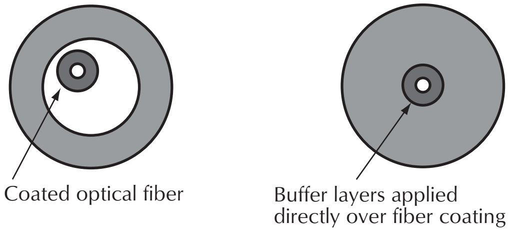
Dispersion. Dispersion is the spreading of a light pulse as it travels down the length of an optical fiber. Dispersion limits the bandwidth or information-carrying capacity of a fiber. In a digital modulated system, this causes the received pulse to be spread out in time. No power is actually lost due to dispersion, but the peak power is reduced as shown in Fig. 15-10. Dispersion can be canceled to zero in single-mode fibers but with multimode it often imposes the system design limit. The units for dispersion are generally given in ns/km.

Loose Tube and Tight Buffer Fiber Jackets. There are basically two types of fiber jacket protection called *loose tube* and *tight buffer*, Fig. 15-11.



A. Input B. Fiber C. Output

Figure 15-10. Dispersion in an optical fiber.



A. Loose tube. B. Tight buffer.

Figure 15-11. Loose tube and tight buffer fiber jackets.

The loose tube is constructed to contain the fiber in a plastic tube that has an inner diameter much larger than the fiber itself. The plastic loose tube is then filled with a gel substance. This allows the fiber to have less stress from the exterior mechanical forces due to the running or pulling of the cable. In multiple fiber loose tube or single fiber loose tube extra strength members are added to keep the fibers free of stress and to help minimize elongation and contraction. Thus, varying the amount of fibers inside the loose tube, the degree of shrinkage can be controlled due to temperature change. This allows for more consistent attenuation over temperature.

The second type, tight buffer, protects the fiber by a direct extrusion of plastic over the basic fiber coating. These tight buffer cables can withstand much greater crush and impact forces without fiber breakage. While the tight buffer has better crush capabilities and is more flexible, it lacks the better attenuation figure of the loose tube due to temperature variations which cause microbending due to sharp bends and twisting of the cable.

Strength members provide for better tensile load parameters similar to coax or electrical audio cables. An optical fiber doesn't stretch very far before it breaks, so the strength members must employ low elongation at the expected tensile loads.

A common strength member used in fiber optic cables for harsh environments is Kevlar™. Kevlar is the material used in bulletproof vests and has the best performance for optical fiber strength members. These strength members are also referred to as tactical optical fiber. They were first used for military communications and were popularized in Operation Desert-Storm in the Iraq and Kuwait war of 1991. These tactical optical fiber cables are impervious to tanks, trucks and bomb explosions. In today's audio applications involving broadcast sports events and news, tactical optical fiber cables have found a niche.

15.4.3 Signal Loss

15.4.3.1 Fiber Optic Transmission Loss (FOTL)

In addition to physical changes to the light pulse which result from frequency or bandwidth limitations, there are also reductions in level of optical power as the light pulse travels to and through the fiber. This optical power loss, or attenuation, is expressed in dB/km (decibels per kilometer). The major causes of optical attenuation in optical fiber systems are:

1. Optical fiber loss.
2. Microbending loss.
3. Connector loss.
4. Splice loss.
5. Coupling loss.

In the ANSI/IEEE Standard 812-1984 the Definition of Terms Relating to Fiber Optics defines attenuation and attenuation coefficient as follows:

Attenuation. In an optical waveguide, the diminution of average optical power. Note: In optical waveguides, attenuation results from absorption, scattering, and other radiation. Attenuation is generally expressed in decibels (dB). However, attenuation is often used as a synonym for attenuation coefficient, expressed as dB/km. This assumes the attenuation coefficient is invariant with length. Also see—attenuation coefficient; coupling loss; differential mode attenuation; equilibrium mode distribution; extrinsic joint loss; leaky modes; macrobend loss; material scattering; microbend loss; Rayleigh scattering; spectral window; transmission loss; waveguide scattering.

Attenuation Coefficient. The rate of diminution of average optical power with respect to distance along the waveguide. Defined by the equation

$$P(z) = P(0)10^{-\left(\frac{\alpha z}{10}\right)} \tag{15-9}$$

where,

$P(z)$ is the power at distance z along the guide,

$P(0)$ is the power at $z = 0$,

α is the attenuation coefficient in dB/km if z is in km.

From this equation,

$$\alpha z = -10\log\left[\frac{P(z)}{P(0)}\right] \tag{15-10}$$

This assumes that α is independent of z; if otherwise, the definition shall be given in terms of incremental attenuation as

$$P(z) = P(0)10^{-\int_0^z \frac{\alpha(z)dx}{10}} \tag{15-11}$$

or, equivalently,

$$\alpha z = -10\frac{d}{dz}\log\left[\frac{P(z)}{P(0)}\right] \tag{15-12}$$

15.4.3.2 Optical Fiber Loss

Attenuation varies with the wavelength of light. Windows are low-loss regions, where fibers carry light with little attenuation. The first generation of optical fibers operated in the first window, around 820 nm to 850 nm. The second window is the zero-dispersion region of 1300 nm, and the third window is the 1550 nm region. A typical 50/125 graded-index fiber offers attenuation of 4 dB/km at 850 nm and 2.5 dB/km at 1300 nm, a 30% increase in transmission efficiency. Attenuation is very high in the regions of 730 nm, 950 nm, 1250 nm, and 1380 nm; therefore, these regions should be avoided.

Evaluating loss in an optical fiber must be done with respect to the transmitted wavelength. Fig. 15-12 shows a typical attenuation curve for a low-loss multimode fiber. Fig. 15-13 does the same for a single-mode fiber; notice the high loss in the mode-transition region, where the fiber shifts from multimode to single-mode

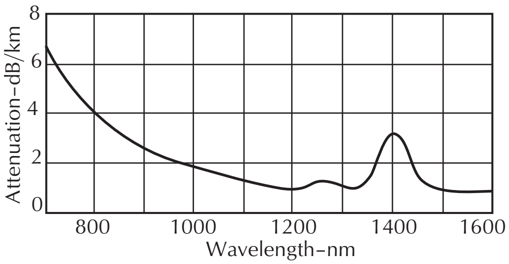


Figure 15-12. Multimode fiber spectral attenuation.

operation. Making the best use of the low-loss properties of the fiber requires that the source emit light in the low-loss regions of the fiber. Plastic fibers are best operated in the visible-light area around 650 nm.

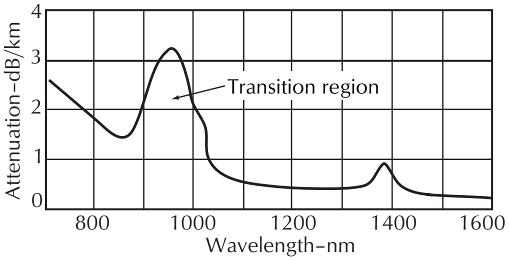


Figure 15-13. Single-mode fiber attenuation.

One important feature of attenuation in an optical fiber is that it is constant at all modulation frequencies within the bandwidth. In a copper cable, attenuation increases with the signal's frequency. The higher the frequency, the greater the attenuation. A 30 MHz signal will be attenuated in a copper cable more than a 15 MHz signal. As a result, signal frequency limits the distance a signal can be sent before a repeater is needed to regenerate the signal. In an optical fiber, both signals will be attenuated the same.

Attenuation in a fiber has three main causes:

1. Scattering.
2. Absorption.
3. Bending (Microbending).

Scattering. Scattering is the loss of optical energy due to imperfections in the fiber and from the basic structure of the fiber. Scattering does just what the term implies: it scatters the light in all directions. The light is no longer directional.

Rayleigh scattering is the same phenomenon that causes a red sky at sunset. The shorter blue wavelengths are scattered and absorbed while the longer red wavelengths suffer less scattering and reach our eyes, so we see a red sunset. Rayleigh scattering comes from density and compositional variations in a fiber that are natural byproducts of manufacturing. Ideally, pure glass has a perfect molecular structure and, therefore, uniform density throughout. In real glass, the density of the glass is not perfectly uniform. The result is scattering.

Since scattering is inversely proportional to the fourth power of the wavelength $(1/\lambda)^4$, it decreases rapidly at longer wavelengths. Scattering represents the theoretical lower limits of attenuation, which are as follows:

- 2.5 dB at 820 nm
- 0.24 dB at 1300 nm
- 0.012 dB at 1550 nm

Absorption. Absorption is the process by which impurities in the fiber absorb optical energy and dissipate it as a small amount of heat. The light becomes dimmer. The high-loss regions of a fiber result from water bands, (where hydroxyl molecules significantly absorb light). Other impurities causing absorption include ions of iron, copper, cobalt, vanadium, and chromium. To maintain low losses, manufacturers must hold these ions to less than one part per billion. Fortunately, modern manufacturing techniques, including making fibers in a very clean environment, permit control of impurities to the point that absorption is not nearly as significant as it was a few years ago.

Microbend Loss. Microbend loss is that loss resulting from microbends, which are small variations or bumps in the core to cladding interface. As shown in Fig. 15-14, microbends can cause high-order modes to reflect at angles that will not allow further reflection. The light is lost.

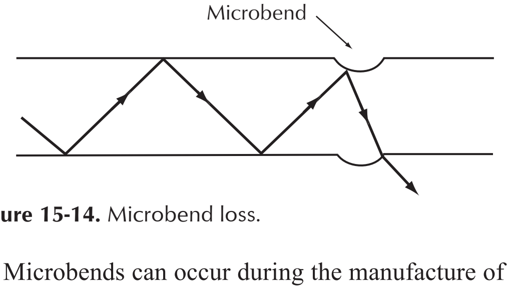


Figure 15-14. Microbend loss.

Microbends can occur during the manufacture of the fiber, or they can be caused by the cable. Manufacturing and cabling techniques have advanced to minimize microbends and their effects.

New Reduced Bend Radius Fibers. Fiber optic cable manufacturers have now significantly reduced the bend radius of the fiber. The reduced bend radius allows for more flexibility allowing installers to bend the fiber around tight corners without any discernible increase in the fiber's attenuation. There are several names given to these optical fibers such as *bend insensitive* or *bend resistant* that can be somewhat misleading when it comes to the selection of the fiber. The user may tend to believe that the reduction of the bend radius will also eliminate any mishandling, temperature extremes, improper routing, or other external forces on the fiber. However, the user should be aware that these factors may not always be true. Selecting a reduced-bend radius-fiber really achieves the improvements of bending the fibers for tighter bends in fiber panels, frames and routing pathways like conduits, raceways and risers.

There is a common basic rule of thumb that the maximum bend radius should be ten times the outside diameter of the cable or approximately 1.5 inches, whichever is greater. This reduced bend radius of the fiber decreases the standard by about 50%, or to 15 mm, without changing the fiber's attenuation.

There have been fiber demonstrations showing a reduced bend radius fiber patch cord and tying a tight knot within the patch cord. Then the patch cord was tested with the tight knot and revealed that no light escaped and also no increase of attenuation was present. These improvements for patch cords have been tremendous, but when it comes to using reduced bend radius for other applications such as in routing in higher densities or easy connector access they will become more critical. Thus, always consult with the manufacturer's guidelines and specifications when selecting reduced bend radius fibers.

Connector Loss. Connector loss is a function of the physical alignment of one fiber core to another fiber core. Scratches and dirt can also contaminate connector surfaces and severely reduce system performance, but most often the connector loss is due to misalignment or end separation.

Several styles of fiber optic connectors are available from major connector suppliers. Typically, each manufacturer has its own design and is generally not compatible with those of other manufacturers. However, things are constantly changing for the better so now all SMA- and ST-type connectors are compatible.

Depending on connector type, different terminating techniques are used:

- Epoxy and Polish—the fiber is epoxied in place in an alignment sleeve, then polished at the ferrule face.
- Optical and Mechanical—both lenses and rigid alignment tubes are commonly used. In addition, index matching mediums may be employed.

The optical power loss of a connector-to-connector interface typically runs between 0.1 dB and 2 dB, depending on the style of the connector and the quality of the preparation.

Splice Loss. Two fibers may be joined in a permanent fashion by fusion, welding, chemical bonding, or mechanical joining. A splice loss that is introduced to the system may vary from as little as 0.01 dB to 0.5 dB.

Coupling Loss. Loss between the fiber and the signal source or signal receiver is a function of both the device and the type of fiber used. For example, LEDs emit light in a broad spectral pattern when compared to laser diodes. Therefore, LEDs will couple more light when a larger core fiber is used, while lasers can be effective with smaller core diameters such as in single-mode systems.

Fiber core size is, therefore, a major factor in determining how much light can be collected by the fiber. Coupled optical power increases as a function of the square of the fiber core diameter.

The *numerical aperture* (NA) is the light gathering ability of a fiber. Only light injected into the fiber at angles greater than the critical angle will be propagated. The material NA relates to the refractive indices of the core and cladding

$$NA = \sqrt{n_1^2 - n_2^2} \tag{15-13}$$

where,
NA is a unitless dimension.

We can also define the angles at which rays will be propagated by the fiber. These angles form a cone, called the acceptance cone, that gives the maximum angle of light acceptance. The acceptance cone is related to the *NA*

$$\begin{aligned} \theta &= \sin^{-1}(NA) \\ NA &= \sin\theta \end{aligned} \tag{15-14}$$

where,
θ is the half-angle of acceptance, Fig. 15-15.

The NA of a fiber is important because it gives an indication of how the fiber accepts and propagates light. A fiber with a large NA accepts light well; a fiber with a low NA requires highly directional light.

In general, fibers with a high bandwidth have a lower NA; thus, they allow fewer modes. Fewer modes mean less dispersion and, hence, greater bandwidth. NAs range from about 0.50 for plastic fibers to 0.21 for

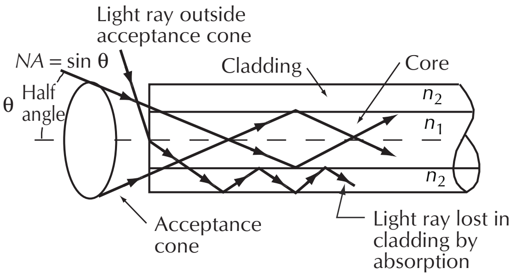

Light ray outside acceptance cone
NA = sin θ
Half angle
θ
Cladding
Core
n_2
n_1
n_2
Acceptance cone
Light ray lost in cladding by absorption

Figure 15-15. Numerical aperture (NA).

graded-index fibers. A large NA promotes more modal dispersion, since more paths for the rays are provided.

Sources and detectors also have an NA. The NA of the source defines the angles of the exiting light. The NA of the detector defines the angles of light that will operate the detector. Especially for sources, it is important to match the NA of the source to the NA of the fiber so that all the light emitted by the source is coupled into the fiber and propagated. Mismatches in NA are sources of loss when light is coupled from a lower NA to a higher one.

15.4.3.3 Attenuation Measurement

In an optical fiber, attenuation measurements require comparison of input and output power P_{in} and P_{out}, respectively. It is measured in decibels as

$$L_{FOP} = -10\log\left(\frac{P_{out}}{P_{in}}\right)\text{in dB} \tag{15-15}$$

where,

the negative sign is added to give attenuation a positive value because the output power is always less than the input power for passive devices,

L_{FOP} is the level of fiber optic power expressed in dB.

Remember these are optical powers, and they are dependent on the wavelength. Optical power digital meters make their measurements readings in either dB or dBm, and also display the wavelength. The optical power level L_{OP} is computed with the equation

$$L_{OP} = 10\log\left(\frac{P_s}{P_r}\right)\text{in dB} \tag{15-16}$$

where,

P_s is the power of the signal,

P_r is the reference power.

If the reference power is 1 mW, then the equation for the optical power level L_{OP} becomes

$$L_{OP} = 10\log\left(\frac{P_s}{1mW}\right)\text{in dB} \tag{15-17}$$

Notice when we know the reference power is 1 mW the unit of level changes to dBm. When the reference power is not specified the unit of level is in dB.

Precise fiber attenuation measurements are based on the *cut-back method test* shown in Fig. 15-16. Here a light source is used to put a signal into the optical fiber; a mode filter is used in graded index or multimode fiber to establish a consistent launch condition to allow consistency of measurements. Although modal conditioning (using mode filter) is beyond the scope of this discussion, it is a very important topic for making measurements in multimode fiber because of the effects of modal conditioning on the values one will measure in this test. Measure the amount of light that comes out at the far end, then cut the fiber back (about 1 to 2 meters) to just past the mode filter. Measure the amount of light that comes out the new end. The difference in the light at one end and that at the other end divided by the length of the fiber gives you the loss per unit length, or the attenuation of the fiber. This is the method used by all manufacturers for testing their fiber.

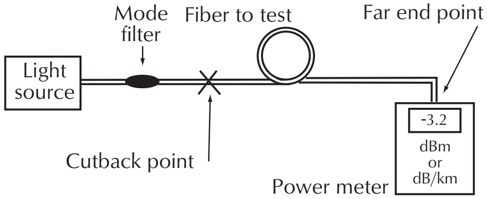


Figure 15-16. The cut-back method for fiber attenuation.

Be aware, however, that this will not accurately measure the loss that light will experience in short multimode fibers because that loss depends on propagation of high-order modes that are eliminated from measurements by adding a mode filter.

Similar measurements can be made on fiber optic cables with mounted connectors by replacing the short cut-back fiber segment with a short jumper cable (including a mode filter if desired). That approach simplifies measurements by avoiding the need to cut fibers at a modest sacrifice in accuracy. One special problem with single-mode fibers is that light can propagate short distances in the cladding, throwing off measurement results by systematically underestimating input coupling losses. To measure true single-mode transmission and coupling, fiber lengths should be at least 20 to 30 m (65 to 100 ft).

Testing fiber optic continuity is important for system function checks. This test for continuity is simple and doesn't require elaborate equipment. A technician on one end shines a flashlight into the fiber, and the technician on the other end looks to see if any light emerges. That quick test can be checked by measuring cable attenuation. Sites of discontinuities can be located with optical time domain reflectometers (OTDRs), as well as attenuation measurements of the cable.

The OTDR contains a high-power laser source and a sensitive photodetector coupled to a signal amplifier that has a wide dynamic range. The output signal is displayed on an integral oscilloscope. OTDRs use the fundamental reflection or backscatter properties of optical fibers by launching a well-defined optical pulse shape into the fiber and measuring and displaying the return level. However, OTDRs are more elaborate and expensive. The alternative for the audio engineer might be an optical fault finder like the one by Tektronix® (Model TOP300) in Fig. 15-17. The TOP300 is a handheld unit which weighs about one pound and incorporates easy to read LEDs. No experience is necessary for the user, just push the buttons and read the LEDs. The strong laser light shows where the fault is located. There are other test instruments for fiber optics which is beyond the scope of this chapter.

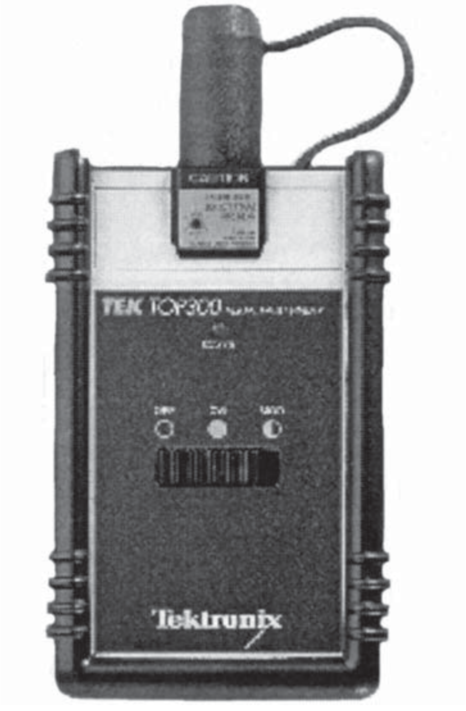


Figure 15-17. An optical fault finder. Courtesy Tektronix®.

15.4.3.4 Advancements in OTDR Testing

The optical time domain reflectometer (OTDR) is designed to troubleshoot fiber breaks and fiber losses. In the past the OTDR was very elaborate and extremely expensive. Fiber optic manufacturers finally have made the OTDR's measurement less complicated. An example of one such device is the OptiFiber® Advanced OTDR by Fluke Networks, Fig. 15-18.

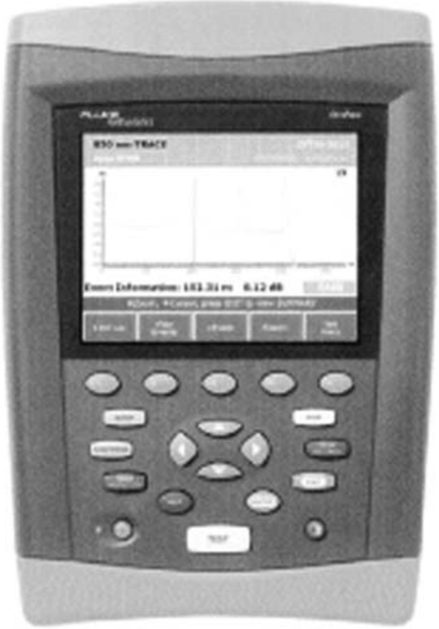


Figure 15-18. OptiFiber Advanced Certifying OTDR. Courtesy of Fluke Networks.

The OptiFiber Advanced OTDR Package will test the fiber link/span, certify it, diagnose it, and document it. This is one of the first certifying OTDRs designed for network owners and installers.

The use of fiber optics in audio and broadcast networks is continually growing, and so are the requirements for testing and certifying. To insure the performance of these optical networks/LANs, network owners are demanding more information that gives them a complete picture of the fiber links. Using this type of OTDR provides a more complete picture.

The OptiFiber is the first test instrument specifically designed to keep network owners and installers on top of the latest requirements for testing and certifying fiber networks. OptiFiber integrates insertion loss and fiber length measurement, OTDR analysis, and fiber connector end face imaging to provide a higher standard of fiber certification and diagnostics. The companion PC software documents, reports, and manages all test data. OptiFiber enables audio network owners of all experience levels to certify fiber to industry and customer specifications, and troubleshoot short-haul connection links and thoroughly document their results.

15.5 Sources

Sources are transmitters of light that can be coupled into fiber optic cables. Basically the two major sources used in fiber optic communications are light emitting diodes (LEDs) and laser diodes. Both are made from semiconductor materials.

LEDs and laser diodes are created from layers of P- and N-type semiconductor materials, forming a junction. Applying a small voltage across the junction causes electrical current, consisting of electrons and holes, to flow. Light photons are emitted from the junction when the electrons and holes combine internal to the junction.

Although the LED provides less power and operates at slower speeds, it is amply suited to applications requiring speeds to several hundred megabits and transmission distances of several kilometers. It is also more reliable, less expensive, has a longer life expectancy, and is easier to use. For higher speeds or longer transmission distances, the laser diode must be considered. Table 15-2 lists the characteristics of typical sources.

Table 15-2. Characteristics of Typical Sources

Type	Output Power (μW)	Peak Wavelength (nm)	Spectral Width (nm)	Rise Time (ns)
LED	250	820	35	12
	700	820	35	6
	1500	820	35	6
LASER	4000	820	4	1
	6000	1300	2	1

Courtesy AMP Incorporated.

15.5.1 LEDs

LEDs are made from a variety of materials located in the Group III to Group V of the Periodic Table of Elements. The color or emission wavelength depends upon the material. Table 15-3 shows some common LED materials used to generate the corresponding colors and wavelengths.

Table 15-3. Materials to Make LEDs and Laser Diodes

Material	Color	Wavelength
Gallium phosphide	green	560 nm
Gallium arsenic phosphide	yellow-red	570–700 nm
Gallium aluminum arsenide	near-infrared	800–900 nm
Indium gallium arsenic phosphide	near-infrared	1300–1500 nm

You might have seen LEDs being used in VU or peak-reading meter displays, or as simple status indicators. LEDs used in fiber optics are designed somewhat differently than a simple display LED. The complexities arise from the desire to construct a source having characteristics compatible with the needs of a fiber optic system. Principal among these characteristics are the wavelength and pattern of emission. There are special packaging techniques for LEDs to couple maximum light output into a fiber, Fig. 15-19.

There are three basic types of designs for fiber optic LEDs:

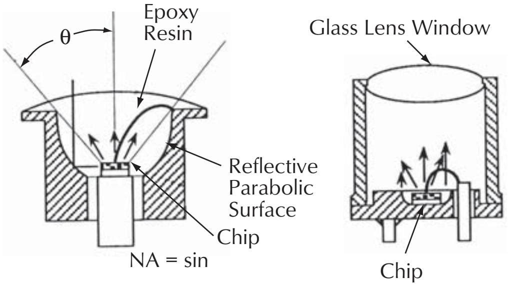
- Surface emitting LED.
- Edge emitting LED.
- Microlensed LED.

Surface Emitting LED. Surface emitting LEDs, Fig. 15-20A, are the easiest and cheapest to manufacture. The result is a low-radiance output whose large emission pattern is not well suited for use with optical fibers. The problem is that only a very small portion of the light emitted can be coupled into the fiber core.

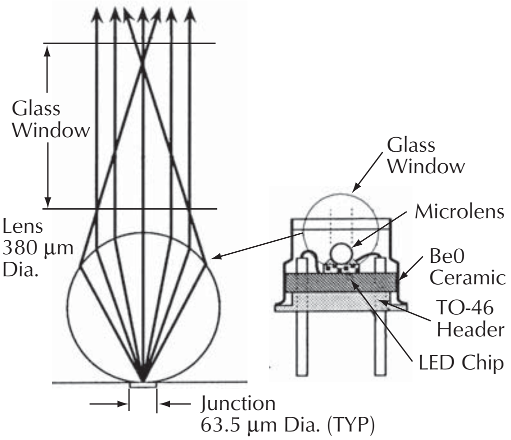
The Burrus LED, named after its inventor Charles A. Burrus of Bell Labs, is a surface-emitting LED with a hole etched to accommodate a light collecting fiber, Fig. 15-21. However, the Burrus LED is not frequently used in modern systems.

Edge Emitting LED. The edge emitting LEDs, Fig. 15-20B, use an active area having stripe geometry. Because the layers above and below the stripe have different refractive indices, carriers are confined by the waveguide effect produced. (The waveguide effect is the same phenomenon that confines and guides the light in the core of an optical fiber.) The width of the emitting area is controlled by etching an opening in the silicon oxide insulating area and depositing metal in the opening. Current through the active area is restricted to the area below the metal film. The result is a high-radiance elliptical output which couples much more light into small fibers than surface emitting LEDs.

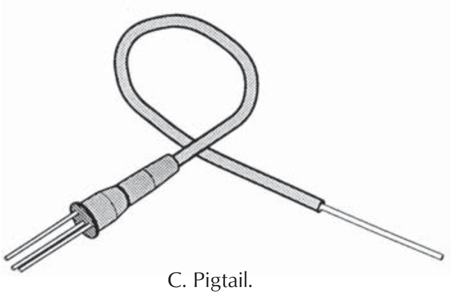
Microlensed LED. More recently, technology has advanced such that it is possible, under production con-


θ
Epoxy Resin
Glass Lens Window
Reflective Parabolic Surface
Chip
NA = sin
Chip

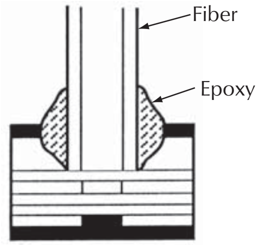
A. Reflecting parabolic surface and glass lens window.


Glass Window
Lens 380 μm Dia.
Junction 63.5 μm Dia. (TYP)
Glass Window
Microlens
Be0 Ceramic
TO-46 Header
LED Chip

B. Microlens.

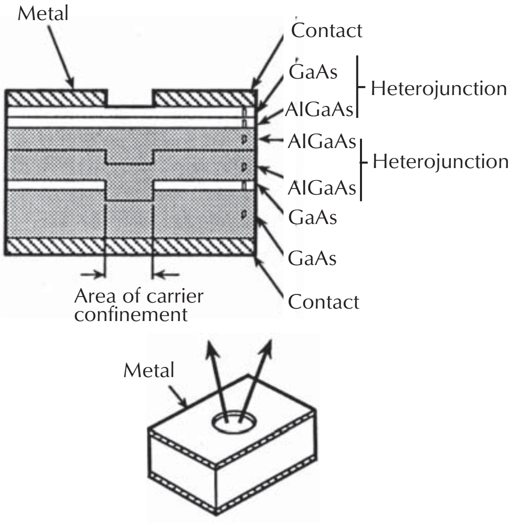


C. Pigtail.

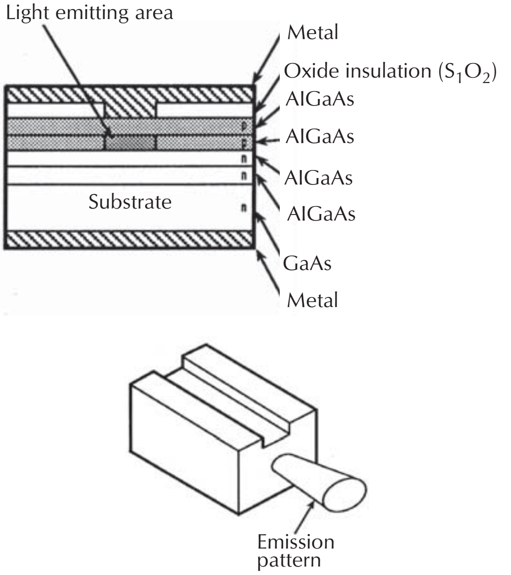

Fiber
Epoxy

D. Burrus diode.

Figure 15-19. Packaging techniques attempt to couple maximum light into a fiber. Courtesy AMP Incorporated.


Metal
Contact
GaAs
AlGaAs
Heterojunction
AlGaAs
Heterojunction
AlGaAs
GaAs
GaAs
Contact
Area of carrier confinement
Metal

A. Surface.


Light emitting area
Metal
Oxide insulation (S_1O_2)
AlGaAs
AlGaAs
AlGaAs
AlGaAs
Substrate
GaAs
Metal
Emission pattern

B. Edge.

Figure 15-20. Surface and edge emitting LEDs. Courtesy AMP Incorporated.

ditions, to place a microscopic glass bead that acts as a lens on top of the diode's microchip structure. This microlensed device has the advantage of direct compatibility with a very wide range of possible fibers. There are also double-lensed versions which allow light to be concentrated into the output fiber pigtail.

Dominant Wavelengths. Most LEDs will have a maximum power at a dominant wavelength lying somewhere within the range of 800 to 850 nm (first window). Some LEDs are available for other wavelengths: either around

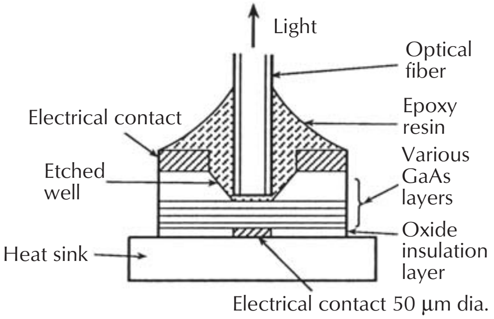


Figure 15-21. The Burrus LED double heterostructure.

1300 nm (second window) or around 1550 nm (third window). The choice is dictated by:

1. Windows—i.e., loss minima, in optical fibers.
2. Availability of suitable detectors.
3. Cost.
4. Minimization of pulse spreading (dispersion) in a fiber.
5. Reliability.

Also the facility for wavelength-division multiplexing (WDM) can also be a factor influencing choice.

15.5.2 Laser Diodes

Laser is an acronym for light amplification by the stimulated emission of radiation. The main difference between an LED and a laser is that the laser has an optical cavity required for lasing, see Fig. 15-22. This cavity, called a Fabry-Perot cavity, is formed by cleaving the opposite end of the chip to form highly parallel, reflective mirrorlike finishes.

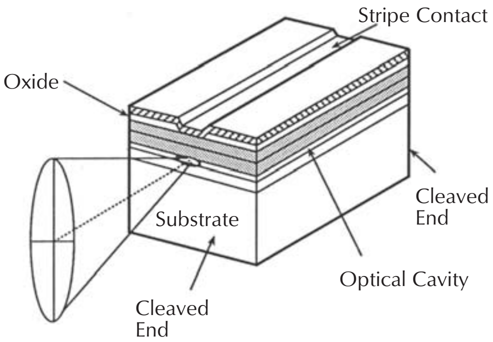


Figure 15-22. Semiconductor laser. Courtesy AMP Incorporated.

At low electrical drive currents, the laser acts like an LED and emits light spontaneously. As the drive current increases, a threshold level is reached, above which lasing action begins. A laser diode relies on high current density (many electrons in the small active area of the chip) to provide lasing action. Some of the photons emitted by the spontaneous action are trapped in the Fabry-Perot cavity, reflecting back and forth from end mirror to end mirror. These photons have an energy level equal to the band gap of the laser material. If one of these photons influences an excited electron, the electron immediately recombines and gives off a photon. Remember that the wavelength of a photon is a measure of its energy. Since the energy of the stimulated photon is equal to the original stimulating photon, its wavelength is equal to that of the original stimulating photon. The photon created is a clone of the first photon. It has the same wavelength, phase, and direction of travel. In other words, the incident photon has stimulated the emission of another photon. Amplification has occurred, and emitted photons have stimulated further emission.

The high drive current in the chip creates population inversion. *Population inversion* is the state in which a high percentage of the atoms move from the ground state to the excited state so that a great number of free electrons and holes exist in the active area around the junction. When population inversion is present, a photon is more likely to stimulate emission than be absorbed. Only above the threshold current does population inversion exist at a level sufficient to allow lasing.

Although some of the photons remain trapped in the cavity, reflecting back and forth and stimulating further emissions, others escape through the two cleaved end faces in an intense beam of light. Since light is coupled into the fiber only from the front face, the rear face is often coated with a reflective material to reduce the amount of light emitted. Light from the rear face can also be used to monitor the output from the front face. Such monitoring can be used to adjust the drive current to maintain constant power level on the output.

Thus, the laser differs from an LED in that laser light has the following attributes:

1. Nearly monochromatic: The light emitted has a narrow band of wavelengths. It is nearly monochromatic, that is, of a single wavelength. In contrast to the LED, laser light is not continuous across the band of its special width. Several distinct wavelengths are emitted on either side of the central wavelength.
2. Coherent: The light wavelengths are in phase, rising and falling through the sine-wave cycle at the same time.

3. Highly directional: The light is emitted in a highly directional pattern with little divergence. Divergence is the spreading of a light beam as it travels from a source.

15.5.3 Superluminescent Diodes (SLDs)

A source called the superluminescent diode (SLD) is now available for use. The performance and cost of the SLD fall somewhere in between the LED and the laser. The SLD was first investigated in 1971 by the Soviet physicist, Kurbatov. The SLD may operate like a edge-emitting LED at low currents, while at high-injection currents, the output power increases superlinearly and the spectral width narrows as a result of the onset of optical gain.

15.5.4 Vertical Cavity Surface Emitting Laser (VCSEL)

A more recent source is the vertical cavity surface emitting laser (VCSEL). It is a specialized laser diode that promises to revolutionize fiber optic communications by improving efficiency and increasing data speed. The acronym VCSEL is pronounced vixel. It is typically used for the 850 nm and 1300 nm windows in fiber optic systems.

15.5.5 LED and Laser Characteristics

15.5.5.1 Output Power

Both LEDs and laser diodes have VI voltage versus current characteristic curves similar to regular silicon diodes. The typical forward voltage drop across LEDs and laser diodes is 1.7 volts.

In general, the output power of sources decreases in the following order: laser diodes, edge emitting LEDs, surface emitting LEDs. Fig. 15-23 shows some curves of relative output power versus input current for LEDs, SLDs, and laser diodes.

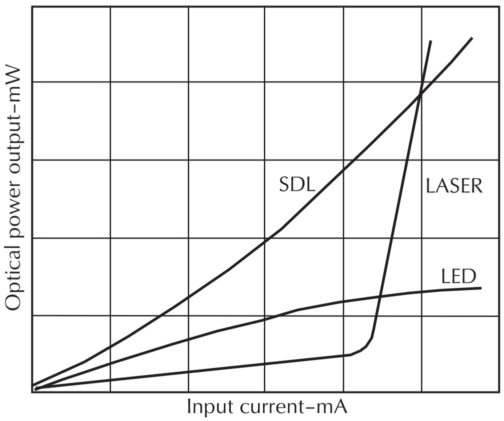


Figure 15-23. Optical power output versus input current for LEDs, SLDs, and laser diodes.

15.5.5.2 Output Pattern

The output or dispersion pattern of light is an important concern in fiber optics. As light leaves the chip, it spreads out. Only a portion of light actually couples into the fiber. A smaller output pattern allows more light to be coupled into the fiber. A good source should have a small emission diameter and a small NA. The emission diameter defines how large the area of emitted light is, and the NA defines at what angles the light is spreading out. If either the emitting diameter or the NA of the source is larger than those of the receiving fiber, some of the optical power will be lost. Fig. 15-24 shows typical emission patterns for the LED, SLD, and laser.

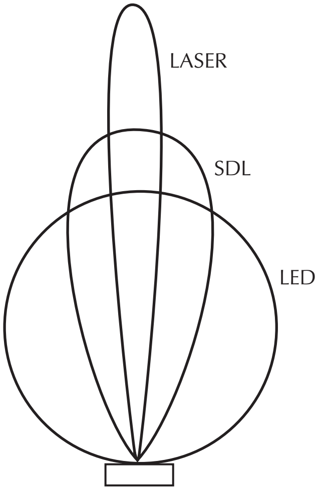


Figure 15-24. Emission patterns of sources.

15.5.5.3 Wavelength

Optical fibers are sensitive to wavelength, therefore the spectral (optical) frequency of a fiber optic source is important. LEDs and laser diodes do not emit a single wavelength; they emit a range of wavelengths. This range is known as the spectral width of the source. It is measured at 50% of the maximum amplitude of the peak wavelength. As an example, if a source has a peak wavelength of 820 nm and a spectral width of 30 nm, its output ranges from 805 nm to 835 nm from the spectral width curve specs. The spectral width of a laser diode is about 0.5 nm to 6 nm; the spectral width of LEDs is much wider—around 20 nm to 60 nm.

15.5.5.4 Speed

A source must turn on and off fast enough to meet the bandwidth requirements of the system. The source speed is specified by rise and fall times. Lasers have rise times of less than 1 ns, whereas LEDs have slower rise times of about 5 ns. A rough approximation of bandwidth for a given rise time is

$$BW = \frac{0.35}{t_r} \tag{15-18}$$

where,

BW is the bandwidth in hertz,

t_r is the rise time in seconds.

15.5.5.5 Lifetime

The expected operating lifetime of a source runs into the millions of hours. Over time, however, the output power decreases due to increasing defects in the device's crystal-line structure. The lifetime of the source is normally considered the time where the peak output power is reduced 50% or 3 dB. In general LEDs have a longer lifetime than laser diodes. As an example, an LED emitting a peak power of 1 mW is considered at the end of its lifetime when its peak power becomes 500 µW or 0.5 mW.

15.5.5.6 Safety

There are a few main precautions to take in the field of fiber optics. Most important is to never look directly into an LED or laser diode! Generally, the light emitted by LEDs is not intense enough to cause eye damage, however, it is best to avoid looking at all collimated beams emitted from LEDs or lasers. Be familiar with the sources used. For more safety information, you can contact the Laser Society of America or OSHA.

15.6 Detectors

The detector performs the opposite function from the source: it converts optical energy to electrical energy. The detector can be called an optoelectronic transducer. The most common detectors in fiber optics are PIN photodiodes, avalanche photodiodes (APD), and integrated detectors-preamplifiers (IDP).

The PIN photodiode is the simplest type of detector, useful for most applications. It is a three-layer semiconductor device having a layer of undoped (or intrinsic) material sandwiched between a layer of positively doped material and negatively doped material. The acronym PIN comes from this ordering: positive, intrinsic, negative. Light falling on the intrinsic layer causes electron-hole pairs to flow as current. In a perfect photodiode, each photon will set an electron-hole pair flowing. In real PIN photodiodes, the conversion from light to electric current is not perfect; only 60% (or less) of the photons reaching the diode causes current flow.

This ratio is the detector's responsitivity. A photodiode has a responsitivity of about 0.6 A/W; in practical terms, an electrical current of 60 µA results for every 100 µW of optical energy striking the diode. Responsivity (R) is the ratio of the diode's output current to input optical power and is given in amperes/watt (A/W). The responsivity also depends on the wavelength of light. Being the simplest device, the PIN photodiode offers no amplification of the signal. Even so, it has several virtues: it is inexpensive, easy to use, and has a fast response time.

The avalanche photodiode (APD) provides some gain and is more sensitive to low-power signals than the PIN photodiode. A photon striking the APD will set a number of electron-hole pairs in motion, which in turn sets other pairs in motion, a phenomenon known as the avalanche effect. A photon initiates an avalanche of current. A typical APD has a responsivity of 15 µA/µW. An additional advantage of the APD is that it is very fast, turning on and off much faster than a photodiode. The drawback to the APD is its complexity and expense. It requires high voltages for operation and is sensitive to variations in temperature. Like the laser as a source, the APD is only used where speeds and distance require it.

The integrated detector-preamplifier (IDP) is a photodetector and transimpedance amplifier in the same integrated circuit. The advantage is that the signal can be amplified or strengthened immediately, before it meets the noise associated with the load resistor. This is important since any following amplifier stages will boost not only the signal but the noise as well. The IDP amplifies the light induced current and provides a usable voltage output. The responsivity of an IDP is in volts/watt (V/W). The responsivity of a typical IDP is about 15 mV/µW. Again, the device has provided gain to overcome noise and provide a suitable SNR.

The characteristics of typical detectors are shown in Table 15-4.

Table 15-4. Characteristics of Typical Detectors

Type	Responsivity	Response Time (ns)
PIN Photodiode	0.5 μA/μW	5
	0.6 μA/μW	1
	0.4 μA/μW	1
APD	75.0 μA/μW	2
	65.0 μA/μW	0.5
IDP	4.5 mV/μW	10
	35.0 mV/μW	35

Courtesy AMP Incorporated.

15.6.1 Quantum Efficiency (η)

Quantum efficiency, another way of expressing a photodiode's sensitivity, is the ratio of photons to the number of electrons set flowing in the external circuit and is expressed either as a dimensionless number or as a percentage. The responsivity can be calculated from the quantum efficiency as follows:

$$R = \frac{\eta q \lambda}{hc} \tag{15-19}$$

where,
q is the charge of an electron,
h is Planck's constant,
c is the velocity of light.

Since q, c, and h are constants, responsivity is simply a function of quantum efficiency and wavelength.

15.6.2 Noise

Several types of noise are associated with the photodetector and with the receiver. Shot noise and thermal noise are particularly important to our understanding of photodiodes in fiber optics.

The noise current produced by a photodiode is called shot noise. Shot noise arises from the discrete nature of electrons flowing across the *p-n* junction of the photodiode. The shot noise can be calculated by using the following equation

$$i_{sn} = \sqrt{2qI_A(BW)} \tag{15-20}$$

where,
q is the charge of an electron (1.6×10^{-19} coulomb),
I_a is the average current (including dark current and signal current),
BW is the receiver bandwidth.

Dark current in a photodiode is the thermally generated current. The term *dark* relates to the absence of light when in an operational circuit.

Thermal noise (i_{tn}), sometimes called Johnson or Nyquist noise, is generated from fluctuations in the load resistance of the photodiode. The following equation can be used to calculate the thermal noise

$$i_{tn} = \sqrt{\frac{4KT(BW)}{R_{eq}}} \tag{15-21}$$

where,
K is Boltzmann's constant (1.38×10^{-23} joules/K),
T is the absolute temperature in Kelvins,
BW is the receiver's bandwidth,
R_{eq} is equivalent resistance, which can be approximated by a load resistor.

Noise in a PIN photodiode is

$$i_n = \sqrt{i_{sn}^2 + i_{TN}^2} \tag{15-22}$$

where,
i_{TN} is the thermal noise.

For an APD, the noise associated with multiplication must also be added.

As a general rule, the optical signal should be twice the noise current to be adequately detected. More optical power may be necessary, however, to obtain the desired SNR.

15.6.3 Bandwidth

The bandwidth, or operating range, of a photodiode can be limited by either its rise time or its RC time constant, whichever results in the slower speed or bandwidth. The bandwidth of a circuit limited by the RC time constant is

$$BW = \frac{1}{2\pi R_{eq} C_d} \tag{15-23}$$

where,
R_{eq} is the equivalent resistance offered by the sum of the load resistance and diode series resistance,
C_d is the diode capacitance including any contribution from the mounting.

A photodiode's response does not completely follow the exponential response of an RC circuit because changes of light frequency or intensity change the parameters. Nevertheless, considering the device equivalent to a low-pass RC filter yields an approximation of its bandwidth. Fig. 15-25 shows the equivalent circuit model of a PIN photodetector.

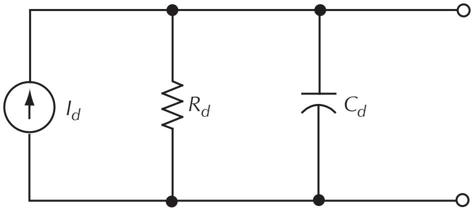


Figure 15-25. Equivalent circuit model of a PIN diode.

15.7 Transmitter/Receiver Modules

In most cases, fiber optic engineers will not design their own transmitters and receivers. They will use completed transmitter-receiver modules. A transmitter module may consist of the following elements:

1. Electronic interface analog/digital input.
2. Analog/digital converter.
3. Drive circuits (preamplifiers, etc.).
4. Optical monitoring circuit.
5. Temperature sensing and control for laser diodes.
6. LED or laser diode as light source.
7. FO connector or pigtail at output.

A receiver module may consist of the following elements:

1. PIN or APD photodiode at the input.
2. Amplification circuits.
3. Signal processor A/D.
4. Analog/digital electrical signal at the output.

Usually, the FO engineer will use a matched pair of transmitter and receiver modules as shown in Fig. 15-26. When considering transmitter/receiver modules one must consider the following requirements:

1. Type of modulation.
2. Bandwidth.
3. Noise.
4. Dynamic range.
5. Electrical and optical interface.
6. Space and cost.

15.8 Transceivers and Repeaters

Transceivers and repeaters are two important components in fiber optics. A transceiver is a transmitter and receiver both in one package to allow transmission and reception from either station. A repeater is receiver driving a transmitter. The repeater is used to boost signals when the transmission distance is so great that the signal will be too highly attenuated before it reaches the receiver. The repeater accepts the signal, amplifies and reshapes it, and sends it on its way by retransmitting the rebuilt signal.

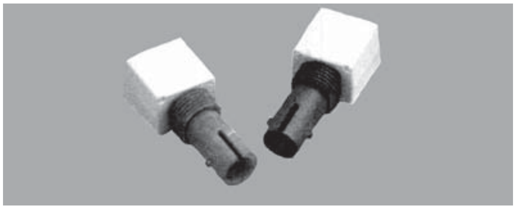


Figure 15-26. A short-wavelength lightwave data link. Courtesy Agilent Technologies.

One advantage of digital transmission is that it uses regenerative repeaters that not only amplify a signal but reshape it to its original form as well. Any pulse distortions from dispersion or other causes are removed. Analog signals use nonregenerative repeaters that amplify the signal, including any noise or distortion. Analog signals cannot be easily reshaped because the repeater does not know what the original signal looked like. For a digital signal, it does know.

15.8.1 Demand on Gigabit Optical Transceivers

The industry is now experiencing that the world needs more bandwidth for today's high-definition technologies. Ethernet has become a standard using both copper and fiber. Manufacturers must keep up with the demand for higher bit rates. The audio/video industry is now employing 1/2/4/10 Gigabit optical transceivers for these high bandwidth applications. Also the need to carry these audio/video signals at distances of 10 km or greater has become a reality. Fiber optics can carry a signal with higher bandwidth and greater distance than their copper counterpart.

The price of copper has gone up tenfold due to the world market consumption of copper, especially in the China market. This has bought the price of fiber and fiber optic transceivers down considerably from 2 years ago. An example is the company 3Com, who manufactures fiber optic transceivers. Most of these fiber optic transceivers employ a SFP (small form factor plug-in) duplex type LC connector. Table 15-5 gives 3Com optical transceiver specification data. Fig. 15-27 is a photo of the 3Com Optical Transceiver.

Table 15-5. 3Com Optical Transceiver—Part No. 3CSFP92

1.25 Gb Gigabit Ethernet/1.063G Fiber Channel	
Application: This 100% 3Com compliant 1000 BASE LX SFP Transceiver is hot-swappable and designed to plug directly into your SFP/GBIC interface slot in your router and switch for Ethernet and Fiber Channel network interface applications.	
Reach	10 km (32, 820 ft)
Fiber Type	SMF (Single Mode Fiber)
Fiber Optic Connector	LC
Center Wavelength λ	1310 nm
Min TX Power	–9.5 dBm
Max Input Power	–3 dBm
RX Sensitivity	–20 dBm
Max Input Power	–3 dBm
Link Budget	10.50 dB
Dimensions	MSA SFP Standard Height: 0.33 inches (8.5 mm) Width: 0.52 inches (13.4 mm) Depth: 2.18 inches (55.5 mm)
Power	3.3V
Operating Temperature	0°C–70°C
Standards:	IEEE 802.3 2003; ANSI X3.297-1997
Compliance	IEC-60825; FDA 21; CFR 1040.10, CFR 1040.11
Warranty	1 Year Full Replacement

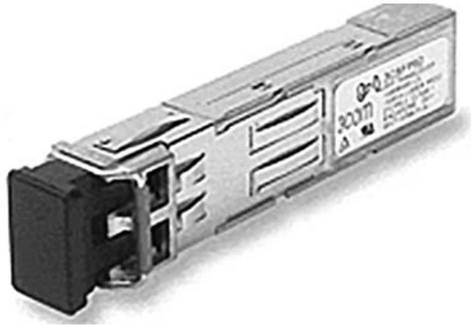


Figure 15-27. 3Com Optical Transceiver. Courtesy of 3Com Corporation.

15.9 The Fiber Optic Link

A basic fiber optic link, as shown in Fig. 15-28, consists of an optical transmitter and receiver connected by a length of fiber optic cable in a point-to-point link. The optical transmitter converts an electrical signal voltage into optical power, which is launched into the fiber by either na LED or laser diode.

At the receiving point, either a PIN or APD photodiode captures the lightwave pulses for conversion back into electrical current.

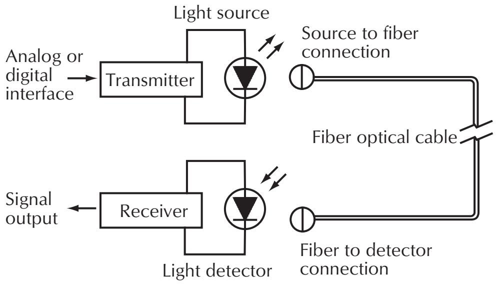

Light source
Source to fiber connection
Analog or digital interface
Transmitter
Fiber optical cable
Signal output
Receiver
Light detector
Fiber to detector connection

Figure 15-28. Basic fiber optic system.

It is the fiber optic system designer's job to determine the most cost- and signal-efficient means to convey this optical power, knowing the trade-offs and limits of various components. He or she must also design the physical layout of the system.

15.9.1 Fiber Optic Link Design Considerations

Fiber optic link design involves power budget analysis and rise time analysis. The power budget calculates total system losses to ensure that the detector receives sufficient power from the source to maintain the required system SNR or bit-error-rate (BER). Rise time analysis ensures that the link meets the bandwidth requirements of the application.

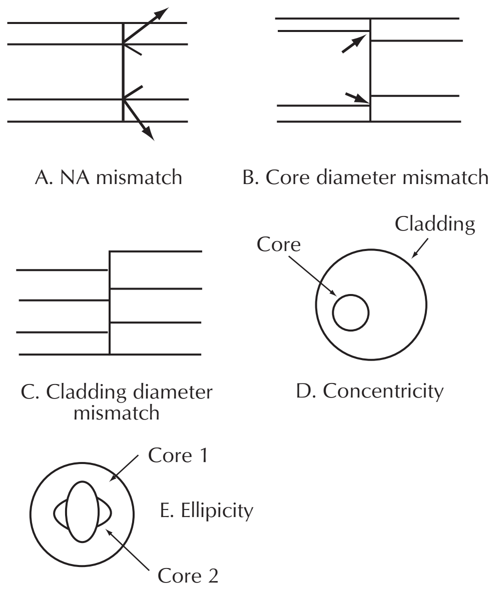
BER is the ratio of correctly transmitted bits to incorrectly transmitted bits. A typical ratio for digital systems is 10^{-9}, which means that one wrong bit is received for every one billion bits transmitted. The BER in a digital system often replaces the SNR in an analog system and is a measure of system quality.

15.9.2 Passive Optical Interconnections

In addition to the fiber, the interconnection system includes the means for connecting the fiber to active devices or to other fibers and hardware for efficiently packaging the system to a particular application. The three most important interconnects are FO connectors, splices, and couplers.

Interconnect losses fall into two categories—intrinsic and extrinsic.

- Intrinsic or fiber-related factors are those caused by variations in the fiber itself, such as NA (numerical aperture) mismatch, cladding mismatches, concentricity, and ellipticity, see Fig. 15-29.
- Extrinsic or connector-related factors are those contributed by the connector itself, Fig. 15-30. The



A. NA mismatch

B. Core diameter mismatch

C. Cladding diameter mismatch

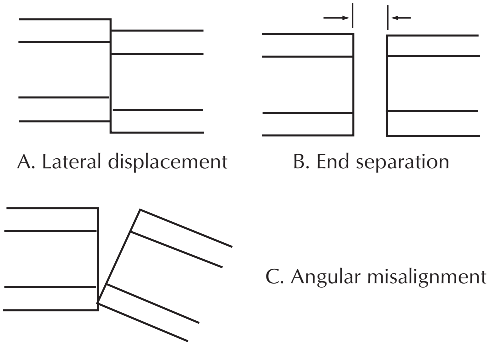
D. Concentricity

E. Ellipicity

Figure 15-29. Intrinsic fiber optic losses.

four main causes of loss that a connector or splice must control are:

1. Lateral displacement.
2. End separation.
3. Angular misalignment.
4. Surface roughness.



A. Lateral displacement

B. End separation

C. Angular misalignment

Figure 15-30. Extrinsic fiber optic losses.

15.9.3 Fiber Optic Connectors

A *fiber optic connector* (FOC) is a device designed to simply and easily permit coupling, decoupling, and recoupling of optical signals or power from each optical fiber in a cable to corresponding fibers in another cable, usually without the use of tools. The connector usually consists of two mateable and demateable parts, one attached to each end of a cable or to a piece of equipment, for the purpose of providing connection and disconnection of fiber optic cables. When selecting FOCs one should look for:

1. Minimum insertion loss.
2. Consistent loss characteristics with little change after many connect/disconnect cycles.
3. Easy installation without expensive tools and special training.
4. Reliability of connection (ruggedness).
5. Low cost.

There are many different types of FOCs being used and newer types are emerging rapidly. We cannot even attempt to cover them all, but will discuss the following popular types in wide use in the communications industry, see Fig. 15-31:

1. Biconic.
2. SMA.
3. FC/PC.
4. ST (preferred for audio applications).
5. SC.
6. D4.
7. FDDI (used in audio for duplex operations).
8. Small form factor connectors:
   LC.
   MT-RJ.

15.9.3.1 Biconic Connector

The biconic connector was invented by AT&T Bell Laboratories. The latest in precision molding techniques are incorporated to yield fractional dB losses. It employs a conic ferrule and has a precision taper on one end that mates to a free-floating precision molded alignment sleeve within the coupling adaptor. While the biconic is still around, it has lost its popularity for the most part.

15.9.3.2 SMA Connector

The SMA was developed by Amphenol Corporation and is the oldest of all FOCs. It is similar to the SMA connector used for microwave applications. The SMA employs a ceramic ferrule and requires preparation of the fiber end for mounting. There are different versions of the SMA by other manufacturers called FSMA.

15.9.3.3 FC/PC Connector

The FC was developed by NTT Nippon Telegraph and Telephone Corp. It has a flat endface on the ferrule that

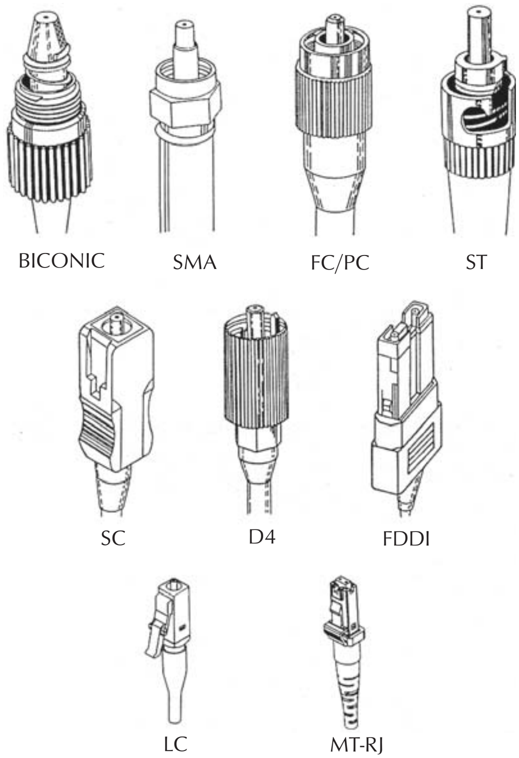


Figure 15-31. Popular fiber optic connectors. (Drawn by Ronald G. Ajemian.)

provides face contact between joining connectors. A modified version of the FC called FC/PC was the first to use physical contact between fiber ends to reduce insertion loss and to increase return loss by minimizing reflections at the interfaces of the joining fibers.

15.9.3.4 ST Connector

The ST (straight through) connector was introduced in early 1985 by AT&T Bell Laboratories. The ST connector design utilizes a spring-loaded twist and lock coupling, similar to the BNC connectors used with coax cable. The ST prevents the fiber from rotating during multiple connections. This insures more consistent insertion loss during multiple connections. The ST is becoming the most popular FOC at the present time because of performance and compatibility. There are many versions of this ST-type connector being offered by other FOC manufacturers, even some that require no epoxy, just a simple crimp.

15.9.3.5 SC Connector

Most recent on the market is the SC-type connector developed by NTT of Japan. It is a molded plastic connector which employs a rectangular cross section and push-pull instead of threaded coupling. The SC achieves a much lower insertion loss than other types of FOCs and has a greater packing density which is useful in multicable installations. Recently Hirose Electric Co. and Seiko of Japan are manufacturing an SC type that has push-pull locking and employs a zirconia ferrule.

15.9.3.6 D4 Connector

The D4 connector was designed by NEC Nippon Electric Corp., Tokyo, Japan. It is similar to the D3, which was a forerunner of the FC.

15.9.3.7 FDDI Connector

The FDDI (Fiber Data Distributed Interface) connector is another recently developed connector. This connector is described and endorsed by the FDDI standard. The IEEE 802.8 (Institute of Electrical and Electronic Engineers) committee now recommends the FDDI connector for all networks involving duplex fiber operation. However, the increasing gain of the duplex SC connector is making it more popular.

15.9.3.8 SFF (Small Form Factor) Connectors

The SFF connectors are fiber optic connectors newly designed to allow for fast, lower cost and increased density of the patch panel/cross-connect field. They are approximately half the size of the traditional ST and SC connectors.

15.9.3.9 LC Connector

The LC SFF connector by Lucent Technologies was introduced to the market in late 1996. The LC connector employs a 1.25 mm ceramic ferrule with a push-pull insertion release mechanism similar to the familiar RJ-45 telephone modular plug. The LC incorporates an antisnag latch that improves durability and reduces cross-connect rearrangement effort. The LC is available in both simplex and duplex types.

15.9.3.10 MT-RJ Connector

The MT-RJ SSF connector was designed by AMP, Inc. (now TYCO) and uses the familiar RJ latching mechanism found in copper systems, but the MT-RJ latch is inherently snag-proof. The single ferrule design of the MT-RJ connector reduces the time and complication of assembly by enabling two-fiber terminations simultaneously.

15.9.3.11 OpticalCon®

Most recent on the market is the OpticalCon® connector developed and introduced in 2006 by Neutrik AG. The OpticalCon® fiber optic connection system consists of a ruggedized all-metal, dust- and dirt-protected chassis and cable connector to increase the reliability. The system is based on a standard optical LC-Duplex connection; however, the OpticalCon® improves this original design to ensure a safe and rugged connection. Due to the compatibility with conventional LC connectors, it offers the choice of utilizing a cost-effective LC connector as a permanent connection, or Neutrik's rugged OpticalCon® cable connector for mobile applications, Fig. 15-32.

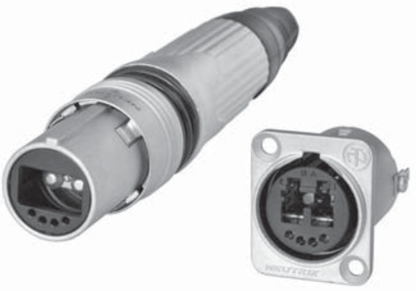


Figure 15-32. OpticalCon® connector. Courtesy of Neutrik AG.

15.9.3.12 Toslink

The Toslink connector was developed by Toshiba of Japan in 1983 and is a registered trademark. This connector was originally designed for a plastic optical fiber of 1 mm diameter. The actual connector/adapter is of a square construction with newer types having a protective flip cap to close the connector adapter when no plug is mated. Also this connector is referred to as JIS FO5 (JIS C5974-1993 FO5) in a simplex type and JIS FO7 for the duplex version, Fig. 15-33.

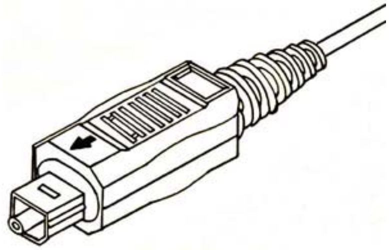


Figure 15-33. Toslink connector.

15.9.3.13 Military Grade Connectors

There are some military-grade types of FOCs that are in use for pro audio that may or may not employ a lens system. These military-grade FOCs go beyond the scope of this chapter.

15.9.4 Connector Installation

The procedure to install a fiber optic connector is similar to that of an electrical connector. However, FOCs require more care, special tools, and a little more time. But as one gains more experience, the time is significantly reduced. The following are steps in making a fiber optic connection:

1. Open the cable.
2. Remove jacketing and buffer layers to expose the fiber.
3. Cut or break the fiber.
4. Insert the fiber into the connector.
5. Attach connector to fiber with epoxy or crimp.
6. Polish or smooth the fiber end.
7. Inspect the fiber ends with a microscope.
8. Seal the connector and fiber cable.

There are presently some FOCs that do not require epoxy or polishing. Things are constantly improving for the better.

15.9.4.1 Current Fiber Optic Connectors for Glass Optical Fibers (GOF)

The LC connector is becoming a de facto standard for pro-audio and video applications. Audio equipment manufacturers are now seeing the benefits of this connector along with its harsh environment type by Neutrik called OpticalCon®. Ongoing work is still in progress for fiber optic cables and connectors by the AES standard group on fiber optic connectors and cables.

15.9.4.2 LC Versus Other Types of Fiber Optic Connectors (ST, SC, etc.)

The LC-type connector over other types of connectors is made for more consistent performance and reliability. The benefits of using the LC type connector are:

1. It is square in shape and keyed, allowing for anti-rotation, which in turn increases the life expectancy of the connector when mated frequently.
2. It provides for quicker access to patch panel applications where the ST connector (for example) has to be turned to lock.
3. It uses a push-pull insertion release mechanism similar to the familiar RJ-45 telephone plug.
4. It allows tightly spaced patch panels, because it need not be turned to be engaged or disengaged.
5. The LC is called a small form factor (SFF) connector, which is about half the size of the SC connector and provides for high-density patch panels.
6. It offers better axial load and side pull features than the ST connector, thus eliminating disturbances caused by the user touching the cable or boot.
7. Users feel comfortable with LC because of its operational resemblance to an RJ-45 electrical connector.
8. The LC type is universally available throughout the world.
9. It eliminates optical discontinuities resulting from pulling on the cable.
10. It is cost effective.

NOTE: For manufacturers with a large base of existing ST or SC type connectors installed, there are hybrid adapters to mate ST or SC connectors to an LC connector, or vice versa, if needed.

15.9.4.3 Fiber Optic Connector Termination Advancements

Fiber connector manufacturers have now improved the termination process in putting a connector together in a few easy steps. One such device is made by Corning, Inc., called the UniCam® connector system. The UniCam® can be best described as a mini pigtail. It incorporates a factory-installed fiber stub that is fully bonded into the connector's ferrule. The other end is precisely cleaved and placed into the patented alignment mechanism of Corning's mechanical splice. Both the field fiber and fiber stub are fully protected from environmental factors. Unlike other no-epoxy, field-installable connectors, the UniCam® connector requires no polishing, which cuts down the time and cost to install and terminate fiber optic connectors. Now it takes about 1 minute to terminate a fiber optic LC, ST, or SC, which is much faster than the time to solder an XLR audio connector. Fig.15-34 shows a UniCam® connector system tool.

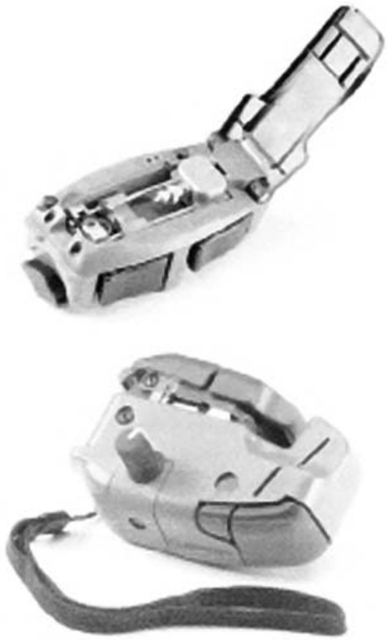


Figure 15-34. UniCam connector system. Courtesy of Corning, Inc.

15.9.4.4 Fiber Optic Splices

A splice is two fibers joined in a permanent fashion by fusion, welding, chemical bonding, or mechanical joining. The three main concerns in a splice are:

1. Splicing loss.
2. Physical durability of the splice.
3. Simplicity of making the splice.

The losses in a fiber optic splice are the same as for FOCs, intrinsic and extrinsic. However the tolerances for a splice are much tighter; therefore, lower attenuation figures are produced.

There are far too many splicing types available to mention; therefore, the following discussion is on splices useful for audio applications. One type by Lucent is called the *CSL LightSplice System*. It provides a fast, easy cleave/mechanical splice for permanent and restoration splicing of single mode and multimode fibers. The CSL LightSplice System features low loss and reflection and unlimited shelf life, and it does not require polishing or the use of adhesives. The splice, Fig. 15-35, also enables the user to visually verify the splicing process.

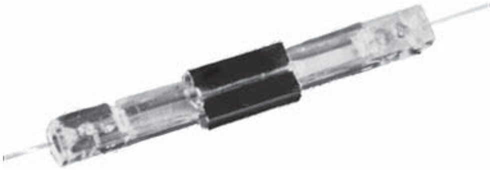


Figure 15-35. Lucent CSL LightSplice. Courtesy Lucent Technologies.

Another splice type is Fibrlok™ optical fiber splice by 3M TelComm Products Division. After cable preparation, the two fibers are inserted into a Fibrlok splice. The assembly tool is then used to close the cap, which forces the three locating surfaces against the fibers. This aligns the fibers precisely and permanently clamps them inside the splice. Fibrlok is for both single- or multimode fibers. The splice, Fig. 15-36, can be performed in about 30 seconds after preparing the fiber ends.

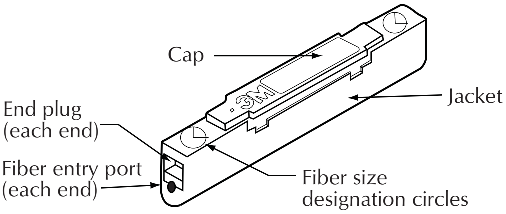


Figure 15-36. Fibrlok optical fiber splice.

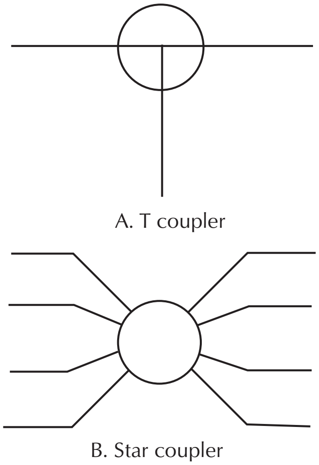
15.10 Couplers

A fiber optic coupler is used to connect three or more fibers together. The coupler is different from a connector or splice which joins only two entities. The fiber optic coupler is more important in fiber optics than for electrical signal transmission because the way optical fibers transmit light makes it difficult to connect more than two points. Fiber optic couplers or splitters were designed to solve that problem.

There are five main concerns when selecting a coupler:

1. Type of fiber used (single- or multimode).
2. Number of input or output ports.
3. Sensitivity to direction.
4. Wavelength selectivity.
5. Cost.

There are two types of passive couplers called the *T* and the *Star* couplers, Fig. 15-37. The T coupler has three ports connected to resemble the letter T. The star coupler can employ multiple input and output ports and the number of inputs can be different from the number of outputs.



A. T coupler

B. Star coupler

Figure 15-37. T and star couplers.

Couplers are quite simple to use. The following calculation must be made:

Excess loss: The losses that are internal to the coupler from scattering, absorption, reflections, misalignment, and poor isolation. Excess loss is the ratio of the sum of all the output power at the output ports to the input power at the input ports. Usually it is expressed in dB.

Insertion loss: This loss is the ratio of the power appearing at a given output port to that of an input port. Thus, insertion loss varies inversely with the number of terminals.

15.11 Fiber Optic System Design

15.11.1 System Specifications

When designing an FO system, it is often best to order all the component parts from one manufacturer; this way you can be sure of the parts being compatible. Many manufacturers have developed complete systems and have components available for the asking. The following are important things to consider when selecting parts and designing a system:

1. If system is analog:
   A. Bandwidth in hertz (Hz) or megahertz (MHz).

B. Distortion in decibels (dB).
C. Operating temperature range in degrees Celsius (°C).

2. If system is digital:
   A. Required BER. Upper BER is usually in megabits per second (Mbps). Lower BER is usually in bits per second (bps).
   B. Operating temperature range in degrees Celsius (°C).
3. If system is audio/video:
   A. Bandwidth in hertz (Hz) or megahertz (MHz).
   B. Distortion in decibels (dB).
   C. Crosstalk in decibels (dB) (for multiple channels).
   D. Operating temperature in degrees Celsius (°C).

15.11.1.1 Transmitter Specifications

1. Input impedance in ohms (Ω).
2. Maximum input signal in dc volts (Vdc), rms or effective volts (Vrms), peak-to-peak volts (Vp-p).
3. Optical wavelength in micrometers (μm) or nanometers (nm).
4. Optical power output in microwatts (μW) or (dBm).
5. Optical output rise time in nanoseconds (ns).
6. Required power supply dc voltage, usually 5 ±0.25 Vdc or 15 ±1 Vdc.

15.11.1.2 Light Source Specifications

1. Continuous forward current in milliamps (mA).
2. Pulsed forward current in milliamps (mA).
3. Peak emission wavelength in nanometers (nm).
4. Spectral width in nanometers (nm).
5. Peak forward voltage in dc volts (Vdc).
6. Reverse voltage in dc volts (Vdc).
7. Operating temperature range in degrees Celsius (°C).
8. Total optical power output in microwatts (μW).
9. Rise/fall times in nanoseconds (ns).

15.11.1.3 Fiber Specifications

1. Mode—single or multimode.
2. Index—step or graded.
3. Attenuation in decibels per kilometer (dB/km).
4. Numerical aperture (NA) (a sine value).
5. Intermodal dispersion in nanoseconds per kilometer (ns/km).
6. Core and cladding diameters in micrometers (μm).
7. Core and cladding index of refraction (a ratio).
8. Bend radius of fiber in centimeters (cm).
9. Tensile strength in pounds per square inch (psi).

15.11.1.4 Cable Specifications

1. Number of fibers (a unit).
2. Core and cladding diameters in micrometers (μm).
3. Cable diameter in millimeters (mm).
4. Minimum bend radius in centimeters (cm).
5. Weight in kilograms per kilometer (kg/km).

15.11.1.5 Detector Specifications

1. Continuous forward current in milliamps (mA).
2. Pulsed forward current in milliamps (mA).
3. Peak reverse voltage in dc volts (Vdc).
4. Capacitance in picofarads (pF).
5. Wavelength in micrometers (μm) or nanometers (nm).
6. Quantum efficiency (η) in percent (%).
7. Responsivity in amps per watt (A/W).
8. Rise/fall time in nanoseconds (ns).
9. Monitor dark current in nanoamperes (nA).
10. Active area diameter in micrometers (μm).
11. Gain coefficient in volts (V) (for APD).
12. Operating temperature in degrees Celsius (°C).

15.11.1.6 Receiver Specifications

1. Output impedance in ohms (Ω).
2. Output signal in dc volts dc (Vdc), rms or effective volts (Vrms), peak-to-peak volts (Vp-p).
3. Optical sensitivity in microwatts (μW), nanowatts (nW), decibels referenced to 1 mW (dBm), or megabits per second (Mbps).
4. Optical wavelength for rated sensitivity in nanometers (nm).
5. Maximum optical input power (peak) in microwatts (μW) or (dBm).
6. Analog/digital rise and fall time in nanoseconds (ns).
7. Propagation delay in nanoseconds (ns).
8. Required power supply in dc volts (Vdc).
9. TTL compatibility.
10. Optical dynamic range in decibels (dB).
11. Operating temperature in degrees Celsius (°C).

15.11.2 Design Considerations

Before designing a fiber optic system, certain factors must be realized.

1. What type of signal information is it?
2. Is signal analog or digital?
3. What is the information bandwidth?
4. What power is required?

5. What is the total length of the fiber optic cable?
6. What is the distance between transmitter and receiver?
7. Are there any physical obstacles that the cable must go through?
8. What are the tolerable signal parameters?
9. What is the acceptable SNR if system is analog?
10. What is the acceptable BER and rise/fall time if system is digital?

Once these parameters are established, the fiber optic system can be designed.

15.11.3 Design Procedures

The procedures for designing a fiber optic system are as follows:

1. Determine the signal bandwidth.
2. If the system is *analog,* determine the SNR. This is the ratio of output signal voltage to noise voltage, the larger the ratio the better. The SNR is expressed in decibels (dB). SNR curves are provided on detector data sheets.
3. If the system is *digital,* determine the BER. A typical good BER is 10^{-9}. BER curves are provided on detector data sheets.
4. Determine the link distance between the transmitter and the receiver.
5. Select a fiber based on attenuation.
6. Calculate the fiber bandwidth for the system. This is done by dividing the bandwidth factor in megahertz per kilometer by the link distance. The bandwidth factor is found on fiber data sheets.
7. Determine the power margin. This is the difference between the light source power output and the receiver sensitivity.
8. Determine the total fiber loss by multiplying the fiber loss in dB/km by the length of the link in kilometers (km).
9. Count the number of FO connectors. Multiply the connector loss (provided by manufacturer data) by the number of connectors.
10. Count the number of splices. Multiply the splice loss (provided by manufacturer data) by the number of splices.
11. Allow 1 dB for source/detector coupling loss.
12. Allow 3 dB for temperature degradation.
13. Allow 3 dB for time degradation.
14. Sum the fiber loss, connector loss, splice loss, source/detector coupling loss, temperature degradation loss, time degradation loss (add values of Steps 8 through 13) to find the total system attenuation.
15. Subtract the total system attenuation from the power margin. If the difference is negative, the light source power receiver sensitivity must be changed to create a larger power margin. A fiber with a lower loss may be chosen or the use of fewer connectors and splices may be an alternative if it is possible to do so without degrading the system.
16. Determine the rise time. To find the total rise time, add the rise time of all critical components, such as the light source, intermodal dispersion, intramodal dispersion, and detector. Square the rise times. Then take the square root of the sum of the total squares and multiply it by a factor of 110%, or 1.1, as in the following equation:

$$\text{System rise time} = 1.1\sqrt{T_1^2 T_2^2 T_3^2 \ldots + T_N^2} \quad (15\text{-}24)$$

15.11.3.1 Fiber Optic System Attenuation

The total attenuation of a fiber optic system is the difference between the power leaving the light source/transmitter and the power entering the detector/receiver. In Fig. 15-38, power entering the fiber is designated as P_S or source power. L_{C1} is the power loss at the source to fiber coupling, usually 1 dB per coupling. The power is of that signal launched into the fiber from the light source at the fiber coupling. L_{F1} represents the loss in the fiber between the source and the splice. Fiber optic cable losses are listed in manufacturer's spec sheets and are in dB/km. L_{SP} represents the power loss at the splice. A typical power loss of a splice is 0.3 to 0.5 dB. L_{F2} represents the power loss in the second length of fiber. L_{C2} is the power loss at the fiber to detector coupling. Finally, P_D is the power transmitted into the detector. Other power losses due to temperature and time degradation are generally around 3 dB loss each. Power at the detector is then generalized as

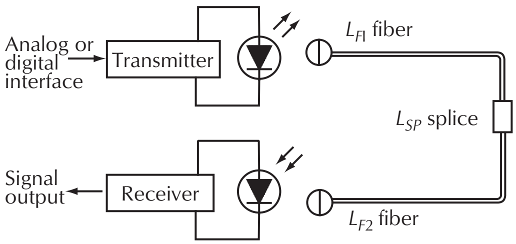
$$P_D = P_S - (L_{C1} + L_{F1} + L_{SP} + L_{F2} + L_{C2})\,. \quad (15\text{-}25)$$

Note: All power and losses are expressed in decibels (dB).

15.11.3.2 Additional Losses

If the core of the receiving fiber is smaller than that of the transmitting fiber, loss is introduced. The following equation can be used to determine the coupling loss from fiber to fiber:

$$L_{dia} = -10\log\left(\frac{dia_r}{dia_t}\right)^2 \quad (15\text{-}26)$$



$P_D = P_S - (L_{C1} + L_{F1} + L_{SP} + L_{F2} + L_{C2})$ dB
P_D = Detector power
P_S = Source power

Figure 15-38. Fiber optic system attenuation.

where,
L_{dia} equals the loss level of the core's diameter,
dia_r is the diameter of the receiving fiber core in μm,
dia_t is the diameter of the transmitting fiber core in μm.

No diameter mismatch loss occurs when light passes from a smaller core to a larger core.

Differences in NA also contribute loss when the input NA of the receiving fiber is less than that of the output NA of the transmitting fiber:

$$L_{NA} = -10\log\left(\frac{NA_r}{NA_t}\right)^2 \tag{15-27}$$

where,
L_{NA} is the loss level of the numerical aperture,
NA_r is the receiving numerical aperture,
NA_t is the transmitting numerical aperture.

Calculation of the NA loss requires that the output NA of the transmitting fiber be known. Since the actual output NA varies with source, fiber length, and modal patterns, using the material NA yields misleading results. No NA mismatch loss occurs when the receiving fiber has an NA greater than that of the transmitting fiber.

The loss of optical power from mismatches in NA and diameter between the source and the core of multimode fiber is as follows:

- When the diameter of the source is greater than the core diameter of the fiber, the mismatch loss is

  $$L_{dia} = -10\log\left(\frac{dia_{fiber}}{dia_{source}}\right)^2 \text{ in dB} \tag{15-28}$$

  where,
  L_{dia} is the level of core diameter mismatch loss.

- No loss occurs when the core diameter of the fiber is larger. When the NA of the source is larger than the NA of the fiber, the mismatch loss is

  $$L_{NA} = -10\log\left(\frac{NA_{fiber}}{NA_{source}}\right)^2 \text{ in dB} \tag{15-29}$$

  where,
  L_{NA} is the numerical aperture mismatch loss.

- No loss occurs when the fiber NA is the larger. Area or diameter loss occurs when a source's area or diameter of emitted light is larger than the core of the fiber. (Area is often used instead of diameter because of the elliptical beam pattern of edge emitters and lasers.) Area or diameter loss is equal to

  $$L_{area} = -10\log\left(\frac{area_{fiber}}{area_{source}}\right) \text{ in dB} \tag{15-30}$$

  where,
  L_{area} is the loss level of the area.

Data sheets for sources often give the area and NA of the output. Although some may not, they may be calculated from information such as polar graphs that are often provided. Calculation of the NA loss and area loss yields an estimate of loss resulting from optical differences between source and fiber. Additional interconnection loss comes from connector related loss, which includes Fresnel reflections and misalignment contributed by a connector.

As with sources, two main causes of loss in coupling light from a fiber into the detector results from mismatches in diameter and NA. When $dia_{det} < dia_{fiber}$, then

$$L_{dia} = -10\log\left(\frac{dia_{det}}{dia_{fiber}}\right)^2 \text{ in dB} \tag{15-31}$$

When $NA_{det} < NA_{fiber}$, then (15-32)

$$L_{NA} = -10\log\left(\frac{NA_{det}}{NA_{fiber}}\right)^2 \text{ in dB} \tag{15-33}$$

where,
L_{dia} is the loss level of the diameter,
L_{NA} is the loss level of the numerical aperture.

Since detectors can be easily manufactured with large active diameters and wide angles of view, such mismatches are less common than with sources. Other losses occur from Fresnel reflections and mechanical misalignment between the connector and the diode package.

15.12 Fiber Optic Considerations

The professional audio engineer, technician or personnel is now facing many new challenges of distributing audio signals. The use of fiber optics is becoming easier, more efficient and cost effective over its copper counterpart. The many breakthroughs in fiber optic technology are leading the way into the future. Glass optical fiber cables are more robust and cost effective enough to use for longer runs exceeding 2 km. Plastic fibers (POFs) are very good at shorter distances (25 ft or less), but they do not meet the fire codes for most building structures. Jitter still seems to be problematic with POFs even at 15 feet in some cases. However, plastic fiber is improving by mixing combinations of glass and plastic which is referred to as plastic-clad silica (PCS) (plastic cladding and glass core). These PCS are being used in industrial applications as well as some telecommunication areas. There are many types of fiber optic system link designs. Usually, the designer is far better off designing and buying components from one or two vendors, which took the guess work out of system compatibility. The advancements of tools for connecting and splicing optical fibers has now become simple and time efficient enough to easily integrate in any audio system. As bandwidths keep increasing, the only thing that will keep up with it is fiber optics. The integrity of the audio signals will not be altered, while keeping the quality at high levels. We are now experiencing fiber to the home FTTH and, to coin a phrase, fiber to the studio FTTS. The audio community is seeing many technological breakthroughs and these fiber optic cables, connectors, and opto-chips are becoming an integral part of pro-audio systems.

15.13 Glossary of Fiber Optic Terms

Absorption: Together with scattering, absorption forms the principal cause of the attenuation of an optical waveguide. It results from unwanted impurities in the waveguide material and has an effect only at certain wavelengths.

Angle of Incidence: The angle between an incident ray and the normal to a reflecting surface.

Attenuation: The reduction of average optical power in an optical waveguide, expressed in dB. The main causes are scattering and absorption, as well as optical losses in connectors and splices. Attenuation or loss is expressed by

$$\mu = -10\log\left(\frac{P_o}{P_i}\right) \text{ dB}$$

Attenuator: An optical element that reduces intensity of a optical signal passing through it (i.e., attenuates it). Example: AT&T makes attenuators built into connectors that incorporate a biconic sleeve consisting of a carbon-coated mylar filter. They come in steps of 6 dB, 12 dB, 16 dB, and 22 dB values.

Avalanche Photodiode (APD): A photodiode designed to take advantage of avalanche multiplication of photocurrent. As the reverse-bias voltage approaches the breakdown voltage, hole-electron pairs created by absorbed photons acquire sufficient energy to create additional hole-electron pairs when they collide with ions; thus, a multiplication or signal gain is achieved.

Axial Ray: A light ray that travels along the axis of an optical fiber.

Backscattering: A small fraction of light that is deflected out of the original direction of propagation by scattering suffers a reversal of direction. In other words, it propagates in the optical waveguide towards the transmitter.

Bandwidth: The lowest frequency at which the magnitude of the waveguide transfer function decreases to 3 dB (optical power) below its zero frequency value. The bandwidth is a function of the length of the waveguide, but may not be directly proportional to the length.

Bandwidth Distance Product (BDP): The bandwidth distance product is a figure of merit that is normalized for a distance of 1 km and is equal to the product of the optical fiber's length and the 3 dB bandwidth of the optical signal. The bandwidth distance product is usually expressed in megahertz*kilometer (MHz*km) or gigahertz*kilometer (GHz*km). For example, a common multimode fiber with bandwidth-distance product of 500 MHz*km could carry a 500 MHz signal for 1 km. Therefore, a 1000 MHz or 1 GHz signal for 0.5 km. Thus, as the distance increases, for 2 km, the BDP would be 250 MHz etc.

Beamsplitter: A device used to divide or split an optical beam into two or more separate beams.

Beamwidth: The distance between two diametrically opposed points at which the irradiance is a specified fraction of the beam's peak irradiance; Beamwidth is most often applied to beams that are circular in cross section.

BER (Bit Error Rate): In digital applications, the ratio of bits received in error to bits sent. *BER*s of one errored bit per billion (1×10^{-9}) sent are typical.

Buffer: Material used to protect optical fiber from physical damage, providing mechanical isolation and/or protection. Fabrication techniques include tight or loose tube buffering, as well as multiple buffer layers.

Burrus LED: A surface-emitting LED with a hole etched to accommodate a light-collecting fiber. Named after its inventor, Charles A. Burrus of Bell Labs.

Chromatic Dispersion: Spreading of a light pulse caused by the difference in refractive indices at different wavelengths.

Cladding: The dielectric material surrounding the core of an optical fiber.

Coarse Wavelength Division Multiplexing (CWDM): CWDM is a cost-effective solution to dense wavelength division modulation (DWDM) that was developed to have channel spacing by the International Telecommunication Union (ITU) in 2002. This standard allows for a 20 nm spacing of channels using wavelengths between 1270 nm and 1610 nm.

Coherent: Light source (laser) in which the amplitude of all waves is exactly equivalent and rise and fall together.

Core: The central region of an optical fiber through which light is transmitted.

Coupler: An optical component used to split or combine optical signals. Also known as a “Splitter,” “T-coupler,” “2 × 2,” or “1 × 2” coupler.

Coupling Loss: The power loss suffered when coupling light from one optical device to another.

Critical Angle: The smallest angle from the fiber axis at which a ray may be totally reflected at the core-cladding interface.

Cutoff Wavelength: The shortest wavelength at which only the fundamental mode of an optical waveguide is capable of propagation.

Dark Current: The external current that, under specified biasing conditions, flows in a photodetector when there is no incident radiation.

Data Rate: The maximum number of bits of information that can be transmitted per second, as in a data transmission link. Typically expressed as megabits per second (Mb/s).

Decibel (dB): The standard unit of level used to express gain or loss of optical or electrical power.

Dense Wavelength Division Multiplexing (DWDM): An enhancement of WDM (see Wavelength Division Multiplexing) that uses many wavelengths in the 1550 nm window (ranges 1530 nm to 1560 nm) for transmitting multiple signals, and often uses fiber optic amplification. Many narrowband transmitters send signals to a DWDM Optical Multiplexer (Mux), which combines all of the signals onto a single fiber. At the other end a DWDM Optical Demultiplexer (Demux) separates the signals out to the many receivers.

Detector: A transducer that provides an electrical output signal in response to an incident optical signal. The current is dependent on the amount of light received and the type of device.

Dispersion: Spread of the signal delay in an optical waveguide. It consists of various components: modal dispersion, material dispersion, and waveguide dispersion. As a result of its dispersion, an optical waveguide acts as a low-pass filter for the transmitted signals.

Ferrule: A component of a fiber optic connection that holds a fiber in place and aids in its alignment.

Fiber Data Distributed Interface (FDDI): An emerging standard developed by AT&T, Hewlett-Packard Co, and Siemens Corp., using a 100 Mbps token ring network that employs dual optical fibers.

Fiber Optic: Any filament or fiber made of dielectric materials, that guides light.

Fiber Optic Link: A fiber optic cable with connectors attached to a transmitter (source) and receiver (detector).

Fresnel Reflection: The reflection of a portion of the light incident on a planar surface between two homogeneous media having different refractive indices. Fresnel reflection occurs at the air–glass interfaces at entrance and exit ends of an optical fiber.

Fundamental Mode: The lowest order mode of a waveguide.

Graded Index Fiber: An optical fiber with a variable refractive index that is a function of the radial distance from the fiber axis.

Incoherent: An LED light source that emits incoherent light as opposed to the laser which emits coherent light. (See Coherent.)

Index Matching Material: A material, often a liquid or cement, whose refractive index is nearly equal to the core index, used to reduce Fresnel reflections from a fiber end face.

Index of Refraction: See Refractive Index.

Injection Laser Diode (ILD): Laser diode.

Insertion Loss: The attenuation caused by the insertion of an optical component. In other words, a connector or coupler in an optical transmission system.

Intensity: Irradiance.

Integrated Optical Components (IOCs): Optical devices (singly or in combination) that use light transmission in waveguides. The waveguides structure and confine the propagating light to a region with one or two very small dimensions of the order of the wavelength of light. A common material used in the fabrication process of an IOC is Lithium Niobate (LiNbO).

Intermodal Distortion: Multimode distortion.

Irradiance: Power density at a surface through which radiation passes at the radiating surface of a light source or at the cross section of an optical waveguide. The normal unit is watts per centimeters squared, or W/cm^2.

Laser Diode (LD): Semiconductor diode that emits coherent light above a threshold current.

Launch Angle: Angle between the propagation direction of the incident light and the optical axis of an optical waveguide.

Launching Fiber: A fiber used in conjunction with a source to excite the modes of another fiber in a particular way. Launching fibers are most often used in test systems to improve the precision of measurements.

Light: In the laser and optical communication fields, the portion of the electromagnetic spectrum that can be handled by the basic optical techniques used for the visible spectrum extending from the near ultraviolet region of approximately 0.3 micron, through the visible region and into the mid infrared region of about 30 microns.

Light Emitting Diode (LED): A semiconductor device that emits incoherent light from a p-n junction when biased with an electrical current in the forward direction. Light may exit from the junction strip edge or from its surface, depending on the device's structure.

Lightwaves: Electromagnetic waves in the region of optical frequencies. The term light was originally restricted to radiation visible to the human eye, with wavelengths between 400 nm and 700 nm. However, it has become customary to refer to radiation in the spectral regions adjacent to visible light (in the near infrared from 700 nm to about 2000 nm) as light to emphasize the physical and technical characteristics they have in common with visible light.

Macrobending: Macroscopic axial deviations of a fiber from a straight line, in contrast to microbending.

Microbending: Curvatures of the fiber that involve axial displacements of a few micrometers and spatial wavelengths of a few millimeters. Microbends cause loss of light and consequently increase the attenuation of the fiber.

Micron: Micrometer (µm). One millionth of a meter (1×10^{-6} m).

Modal Dispersion: Pulse spreading due to multiple light rays traveling different distances and speeds through an optical fiber.

Modal Noise: Disturbance in multimode fibers fed by laser diodes. It occurs when the fibers contain elements with mode-dependent attenuation, such as imperfect splices, and is more severe the better the coherence of the laser light.

Modes: Discrete optical waves that can propagate in optical waveguides. They are eigenvalue solutions to the differential equations that characterize the waveguide. In a single-mode fiber, only one mode, the fundamental mode, can propagate. There are several hundred modes in a multimode fiber that differ in field pattern and propagation velocity. The upper limit to the number of modes is determined by the core diameter and the numerical aperture of the waveguide.

Modified Chemical Vapor Deposition (MCVD) Technique: A process in which deposits are produced by heterogeneous gas/solid and gas/liquid chemical reactions at the surface of a substrate. The MCVD method is often used in fabricating optical waveguide preforms by causing gaseous material to react and deposit glass oxides. Typical starting chemicals include volatile compounds of silicon, germanium, phosphorus, and boron, which form corresponding oxides after heating with oxygen or other gases. Depending on its type, the preform may be processed further in preparation for pulling into an optical fiber.

Monochromatic: Consisting of a single wavelength. In practice, radiation is never perfectly monochromatic but, at best, displays a narrow band of wavelengths.

Multimode Distortion: The signal distortion in an optical waveguide resulting from the superposition of modes with differing delays.

Multimode Fiber: Optical waveguide whose core diameter is large compared with the optical wavelength and in which, consequently, a large number of modes are capable of propagation.

Nanometer (nm): One billionth of a meter (1×10^{-9} m).

Noise Equivalent Power (NEP): The rms value of optical power that is required to produce an rms SNR of 1; an indication of noise level that defines the minimum detectable signal level.

Numerical Aperture: A measure of the range of angles of incident light transmitted through a fiber. Depends on the differences in index of refraction between the core and the cladding.

Optical Fiber Class (OM1, OM2, OM3, and OS1 designations in accordance with ISO11801: Bandwidth and the maximum transmission distance of different optical fiber classes for 10G Ethernet application, Table 15-6.

Optical Time Domain Reflectometer (OTDR): A method for characterizing a fiber wherein an optical pulse is transmitted through the fiber and the resulting backscatter and reflections to the input are measured as a function of time. Useful in estimating the attenuation coefficient as a function of distance and identifying defects and other localized losses.

Optoelectronic: Any device that functions as an electrical-to-optical or optical-to-electrical transducer.

Optoelectronic Integrated Circuits (OEICs): Combination of electronic and optical functions in a single chip.

Peak Wavelength: The wavelength at which the optical power of a source is at a maximum.

Photocurrent: The current that flows through a photosensitive device, such as a photodiode, as the result of exposure to radiant power.

Photodiode: A diode designed to produce photocurrent by absorbing light. Photodiodes are used for the detection of optical power and for the conversion of optical power into electrical power.

Photon: A quantum of electromagnetic energy.

Pigtail: A short length of optical fiber for coupling optical components. It is usually permanently fixed to the components.

PIN-FET Receiver: An optical receiver with a PIN photodiode and low noise amplifier with a high impedance input, whose first stage incorporates a field-effect transistor (FET).

PIN Photodiode: A diode with a large intrinsic region sandwiched between *p*-doped and *n*-doped semiconducting regions. Photons in this region create electron hole pairs that are separated by an electric field thus generating an electric current in the load circuit.

Plastic Optical Fiber (POF): An optical fiber composed of plastic instead of glass. POFs are used for short distances of typically 25 ft or less.

Table 15-6. Bandwidth and the Maximum Transmission Distance of Different Optical Fiber Classes for 10G Ethernet Application

Fiber Type	Bandwidth 850 nm MHz*km	Bandwidth 1300 nm MHz*km	1 Gbps Transmission Distance		10 Gbps Transmission Distance		Fiber Class
Multimode			**@850 nm**	**@1300 nm**	**@850 nm**	**@1300 nm**	
Traditional 62.5/125 µm	200	500	275 m	550 m	33 m	300 m	OM1
Traditional 50/125 µm	400	800	500 m	1000 m	66 m	450 m	OM1
Traditional 50/125/62.5 µm	500	500	550 m	550 m	82 m	300 m	OM2
50/125 µm-110	600	1200	750 m	2000 m	110 m	850 m	OM2+
50/125 µm-150	700	500	750 m	550 m	150 m	300 m	OM2
50/125 µm-300	1500	500	1000 m	550 m	300 m	300 m	OM3
50/125 µm-550	3500	500	1000 m	550 m	550 m	550 m	NA
Single Mode			**@1310 m**	**@1550 nm**	**1310/1383/1550 nm**		
Traditional 9/125 µm			5000 m		10000 m–40000 m		OS1

Preform: A glass structure from which an optical fiber waveguide may be drawn.

Primary Coating: The plastic coating applied directly to the cladding surface of the fiber during manufacture to preserve the integrity of the surface.

Ray: A geometric representation of a light path through an optical medium; a line normal to the wavefront indicating the direction of radiant energy flow.

Rayleigh Scattering: Scattering by refractive index fluctuations (inhomogeneities in material density or composition) that are small with respect to wavelength.

Receiver: A detector and electronic circuitry to change optical signals into electrical signals.

Receiver Sensitivity: The optical power required by a receiver for low error signal transmission. In the case of digital signal transmission, the mean optical power is usually quoted in watts or dBm (decibels referenced to 1 mW).

Reflection: The abrupt change in direction of a light beam at an interface between two dissimilar media so that the light beam returns to the media from which it originated.

Refraction: The bending of a beam of light at an interface between two dissimilar media or in a medium whose refractive index is a continuous function of position (graded index medium).

Refractive Index: The ratio of the velocity of light in a vacuum to that in an optically dense medium.

Repeater: In a lightwave system, an optoelectronic device or module that receives an optical signal, converts it to electrical form, amplifies or reconstructs it, and retransmits it in optical form.

Responsivity: The ratio of detector output to input, usually measured in units of amperes per watt (or microamperes per microwatt).

Single-Mode Fiber: Optical fiber with a small core diameter in which only a single mode—the fundamental mode—is capable of propagation. This type of fiber is particularly suitable for wideband transmission over large distances, since its bandwidth is limited only by chromatic dispersion.

Source: The means (usually LED or laser) used to convert an electrical information carrying signal into a corresponding optical signal for transmission by an optical waveguide.

Splice: A permanent joint between two optical waveguides.

Spontaneous Emission: This occurs when there are too many electrons in the conduction band of a semiconductor. These electrons drop spontaneously into vacant locations in the valence band, a photon being emitted for each electron. The emitted light is incoherent.

ST Connector: A type of connector used on fiber optic cable utilizing a spring-loaded twist and lock coupling similar to the BNC connectors used with coax cable.

Star Coupler: An optical component used to distribute light signals to a multiplicity of output ports. Usually the number of input and output ports are identical.

Step Index Fiber: A fiber having a uniform refractive index within the core and a sharp decrease in refractive index at the core-cladding interface.

Stimulated Emission: A phenomenon that occurs when photons in a semiconductor stimulate available excess charge carriers to the emission of photons. The emitted light is identical in wavelength and phase with the incident coherent light.

Superluminescent Diodes (SLDs): Superluminescent diodes (SLDs) are distinguished from both laser diodes and LEDs in that the emitted light consists of amplified spontaneous emission having a spectrum much narrower than that of LEDs but wider than that of lasers.

T (or Tee) Coupler: A coupler with three ports.

Threshold Current: The driving current above which the amplification of the light-wave in a laser diode becomes greater than the optical losses, so that stimulated emission commences. The threshold current is strongly temperature dependent.

Total Internal Reflection: The total reflection that occurs when light strikes an interface at angles of incidence greater than the critical angle.

Transmission Loss: Total loss encountered in transmission through a system.

Transmitter: A driver and a source used to change electrical signals into optical signals.

Tree Coupler: An optical component used to distribute light signals to a multiplicity of output ports. Usually the number of output ports is greater than the number of input ports.

Vertical Cavity Surface Emitting Laser (VCSEL): A specialized laser diode that promises to revolutionize fiber optic communications by improving efficiency and increasing data speed. The acronym VCSEL is pronounced vixel. Typically used for the 850 nm and 1300 nm windows.

Y Coupler: A variation on the T coupler in which input light is split between two channels (typically planar waveguide) that branch out like a Y from the input.

Wavelength Division Multiplexing (WDM): Simultaneous transmission of several signals in an optical waveguide at differing wavelengths.

Window: Refers to ranges of wavelengths matched to the properties of the optical fiber. The window ranges for fiber optics are the following: First window, 820 nm to 850 nm; second window, 1300 nm to 1310 nm; and third window, 1550 nm.

Bibliography

R. Ajemian, and A. Grundy, “Fiber Optics: The New Medium for Audio,” *J. Audio Eng. Soc.*, Vol. 38, no. 3, March 1990.

R.G. Ajemian, “Fiber Optic Connector Considerations for Professional Audio,” *J. Audio Eng. Soc.*, Vol. 40, no. 6, June 1992.

R.G. Ajemian, “A Tutorial: Fiber Optic Cables and Connectors for Pro-Audio,” *Audio Engineering Society,* 123rd Convention Paper 7210, October 2007.

R.G. Ajemian, *Handbook of Fiber Optic Connectors,* abc Teletraining, Inc. (AVO), 2000.

M.K. Barnoski, *Fundamentals of Optical Fiber Communications,* New York: Academic Press, 1981.

W.B. Bielowski, “Low Loss Waveguide Communication Technology,” Corning, NY: Corning Glass Works.

Corning, Inc., *Just the Facts* magazine, New York.

P. De Lancie, “Optic Link,” *Mix Magazine*, August 1989.

“Guide to Fiber Optics,” Richmond, IN: Belden Wire and Cable Company.

Jeff Hecht, *Understanding Fiber Optics,* Carmel, IN: SAMS, A division of Macmillan Computer Publishing, 1998.

Hentschel, C., *Fiber Optics Handbook,* Palo Alto, CA: Hewlett Packard, 1988.

“IEEE Standard Definitions of Terms Relating to Fiber Optics ANSI/IEEE Std 812-1984,” New York: The Institute of Electrical and Electronics Engineers, Inc.

Gerd Keiser, *Optical Fiber Communications,* New York: McGraw-Hill, 1999.

M. Lambert, “Making The Digital Connection,” *Mix Magazine*, May 1989.

“Lightguide Catalog,” Berkeley Heights, NJ: AT&T, 1990.

Mahlke, Gunther, and Gossing, Peter, *Fiber Optic Cables*, New York: John Wiley & Sons, 1987.

Neutrik USA, http://www.neutrik.com, Product Catalog (OpticalCon), 195 Lehigh Avenue, Lakewood, NJ.

“News of the Week (Scientists create smallest laser),” *Telephony*, July 24, 1989.

“Oak Link Catalog,” Germany: Klotz & Co., 1989.

“Optimate Catalog 86-792,” Harrisburg, PA: AMP Inc., May 1987.

Owl Fiber Optics, http://www.owlfo.com, Fiber Optic Glossary, Flushing, NY

J.C. Palais, *Fiber Optic Communications*, Englewood Cliffs, NJ: Prentice Hall, 1998.

“Practical Testing of Fiber Optic Systems and Components Seminar Workbook,” Boston, MA: Fotec, Inc.

Seippel, R. G., *Optoelectronics for Technology & Engineering,* Englewood Cliffs, NJ: Prentice Hall, 1989.

D. J. Sterling, Jr., *Technician's Guide to Fiber Optics*, New York: Delmar Publishers Inc., 1999.

"Technical Data for Fiber Optic Link Toslink," Burlington, MA: Toshiba America, Inc., 2006.

"Effective Integration of Reduced Bend Radius Fiber into the Network," ADC Telecommunications, Inc., September 2007, Minneapolis, MN 55440-1101.

"Optical Transceiver Installation Guide", November 2003, 3Com Corporation, Southborough, MA USA 01772-1886.

Corning Cable Systems LLC, "UniCam Pretium Connector Catalog", June 2007, Hickory, NC 28603-0489 USA

"OptiFiber® Certifying OTDR", *Fluke Networks,* Everett WA 98203, USA.

This page intentionally left blank

Part 3

Electroacoustic Devices

This page intentionally left blank

Chapter 16

Microphones

by Glen Ballou, Joe Ciaudelli, and Volker Schmitt

16.1 Introduction 493
16.2 Pickup Patterns 493
  16.2.1 Omnidirectional Microphones. 494
  16.2.2 Bidirectional Microphones 495
  16.2.3 Unidirectional Microphones 495
    16.2.3.1 Proximity Effects 497
16.3 Types of Transducers 502
  16.3.1 Carbon Microphones 502
  16.3.2 Crystal and Ceramic Microphones 503
  16.3.3 Dynamic Microphones 503
    16.3.3.1 Unidirectional Microphones 506
  16.3.4 Capacitor Microphones 507
    16.3.4.1 Voltage Divider Capacitor Microphone 508
    16.3.4.2 Phantom Power for Capacitor Microphones 510
    16.3.4.3 Capacitor Radio-Frequency, Frequency- Modulated Microphones 511
  16.3.5 Electret Microphones. 515
16.4 Microphone Sensitivity[7] 516
  16.4.1 Open-Circuit Voltage Sensitivity 516
  16.4.2 Maximum Power Output Sensitivity[7] 516
  16.4.3 Electronic Industries Association (EIA) Output Sensitivity 517
  16.4.4 Various Microphone Sensitivities 518
  16.4.5 Microphone Thermal Noise. 518
16.5 Microphone Practices 519
  16.5.1 Placement 519
    16.5.1.1 Microphone-to-Source Distance 519
    16.5.1.2 Distance from Large Surfaces 519
    16.5.1.3 Behind Objects 520
    16.5.1.4 Above the Source 520
    16.5.1.5 Direct versus Reverberant Field 520
  16.5.2 Grounding 520
  16.5.3 Polarity 521
  16.5.4 Balanced or Unbalanced 521
  16.5.5 Impedance 522
16.6 Miscellaneous Microphones 522
  16.6.1 Pressure Zone Microphones (PZM) 522
    16.6.1.1 Phase Coherent Cardioid (PCC) 524

16.6.1.2 Directivity . . . 524
16.6.2 Lavalier Microphones . . . 527
16.6.3 Head-Worn Microphones . . . 528
16.6.4 Base Station Microphones . . . 530
16.6.5 Differential Noise-Canceling Microphones . . . 530
16.6.6 Controlled-Reluctance Microphones . . . 530
16.6.7 Handheld Entertainer Microphones . . . 530
16.6.8 Pressure-Gradient Condenser Microphones . . . 531
16.6.9 Interference Tube Microphone . . . 532
16.6.10 Rifle Microphones . . . 537
16.6.11 Parabolic Microphones . . . 537
16.6.12 Zoom Microphones . . . 538
16.6.12.1 Distance Factor . . . 539
16.6.12.2 Operation of Zoom Microphones . . . 539
16.6.12.3 Zoom Microphone Video Camera Linkage . . . 540
16.6.13 Automatic Microphone Systems . . . 540
16.6.14 PolarFlex™ Microphone System. . . . 540
16.7 Stereo Microphones . . . 542
16.7.1 Coincident Microphones . . . 542
16.7.2 XY Stereo Technique . . . 543
16.7.3 The ORTF Technique . . . 544
16.7.4 The M/S Stereo Technique . . . 544
16.7.5 The Stereo Boom Microphone . . . 548
16.7.6 SASS Microphones . . . 550
16.7.7 Surround Sound Microphone System . . . 550
16.7.7.1 Schoeps 5.1 Surround System . . . 550
16.7.7.2 Holophone® H2-PRO Surround Sound System . . . 551
16.7.7.3 Holophone H4 SuperMINI Surround Sound System . . . 553
16.8 Microphones for Binaural Recording . . . 553
16.8.1 Artificial Head Systems . . . 553
16.8.2 In the Ear Recording Microphones[18] . . . 555
16.9 USB Microphones . . . 556
16.10 Wireless Communication Systems . . . 557
16.10.1 Criteria for Selecting a Wireless Microphone . . . 558
16.10.1.1 Frequency Bands of Operation . . . 558
16.10.1.2 Adjustment of the System's Operating Frequency . . . 560
16.10.1.3 Capture Ratio and Muting . . . 560
16.10.1.4 RF Power Output and Receiver Sensitivity . . . 560
16.10.1.5 Frequency Response, Dynamic Range, and Distortion . . . 560
16.10.2 Receiving Antenna Systems . . . 561
16.10.2.1 Antenna Placement . . . 562
16.10.3 Companding . . . 563
16.10.4 Waterproof Wireless Microphone Systems . . . 563
16.11 Multichannel Wireless Microphone and Monitoring Systems . . . 564
16.11.1 Introduction . . . 564
16.11.2 Frequencies . . . 564
16.11.3 Spacing . . . 564
16.11.4 Frequency Deviation . . . 564
16.11.5 Frequency Coordination . . . 565
16.11.6 Transmitter Considerations . . . 566

    - 16.11.6.1 Range and RF Power 566
    - 16.11.6.2 Dc-to-Dc Converter 567
    - 16.11.6.3 Audio Processing 567
    - 16.11.6.4 Transmitter Antenna 567
  - 16.11.7 Receiver Considerations 568
    - 16.11.7.1 RF Signal Level 568
  - 16.11.8 Antennas 568
    - 16.11.8.1 Antenna Cables and Related Systems 569
    - 16.11.8.2 Splitter Systems 570
  - 16.11.9 Wireless Monitor Systems. 571
  - 16.11.10 System Planning for Multichannel Wireless Systems. 571
  - 16.11.11 Future Considerations: Digital Wireless Transmission. 572
    - 16.11.11.1 Starting with Sound and the Related Data Rate 572
  - 16.11.12 Conclusion. 573
- 16.12 Microphone Accessories 573
  - 16.12.1 Inline Microphone Processors. 573
    - 16.12.1.1 Sabine Phantom Mic Rider Pro Series 3 573
    - 16.12.1.2 Sabine Phantom Mic Rider Pro Series 2 573
    - 16.12.1.3 Sabine Phantom Mic Rider Series 1 574
    - 16.12.1.4 Lectrosonics UH400A Plug-On Transmitter 574
    - 16.12.1.5 MXL Mic Mate™ USB Adapter 575
  - 16.12.2 Windscreens and Pop Filters. 576
    - 16.12.2.1 Wind Noise Reduction Figures for Rycote Windshielding Devices 576
    - 16.12.2.2 Reflexion Filter 578
  - 16.12.3 Shock Mounts 578
  - 16.12.4 Stands and Booms. 581
  - 16.12.5 Attenuators and Equalizers 581
- 16.13 Microphone Techniques 582
  - 16.13.1 Stereo Micing Techniques. 582
  - 16.13.2 Microphone Choice. 583
  - 16.13.3 Microphone Characteristics. 583
    - 16.13.3.1 Directional Pattern 584
  - 16.13.4 Specific Micing Techniques 584
    - 16.13.4.1 Musicians 584
    - 16.13.4.2 Drums 584
    - 16.13.4.3 Piano 586
    - 16.13.4.4 Vocals 588
    - 16.13.4.5 Singers 588
    - 16.13.4.6 Group Vocals 589
    - 16.13.4.7 Lectern Microphones 589
    - 16.13.4.8 Strings 590
    - 16.13.4.9 Horns 591
    - 16.13.4.10 Woodwinds 591
    - 16.13.4.11 Electric Instruments 592
    - 16.13.4.12 Percussion 592
  - 16.13.5 Conclusion. 593
- Acknowledgments 593
- References 593
- Bibliography. 594

This page intentionally left blank

16.1 Introduction

All sound sources have different characteristics; their waveform varies, their phase characteristics vary, their dynamic range and attack time vary and their frequency response varies, just to name a few. No one microphone will reproduce all of these characteristics equally well. In fact, each sound source will sound better or more natural with one type or brand of microphone than all others. For this reason we have and always will have many types and brands of microphones.

Microphones are electroacoustic devices that convert acoustical energy into electrical energy. All microphones have a diaphragm or moving surface that is excited by the acoustical wave. The corresponding output is an electrical signal that represents the acoustical input.

Microphones fall into two classes: *pressure* and *velocity.* In a *pressure* microphone the diaphragm has only one surface exposed to the sound source so the output corresponds to the instantaneous sound pressure of the impressed sound waves. A pressure microphone is a *zero-order gradient microphone,* and is associated with omni-directional characteristics.

The second class of microphone is the *velocity* microphone, also called a *first-order gradient microphone,* where the effect of the sound wave is the difference or gradient between the sound wave that hits the front and the rear of the diaphragm. The electrical output corresponds to the instantaneous particle velocity in the impressed sound wave. Ribbon microphones as well as pressure microphones that are altered to produce front-to-back discrimination are of the velocity type.

Microphones are also classified by their pickup pattern or how they discriminate between the various directions the sound source comes from, Fig. 16-1. These classifications are:

- *Omnidirectional*—pickup is equal in all directions.
- *Bidirectional*—pickup is equal from the two opposite directions (180°) apart and zero from the two directions that are 90° from the first.
- *Unidirectional*—pickup is from one direction only, the pickup appearing cardioid or heart-shaped.

The air particle relationships of the air particle displacement, velocity, and acceleration that a microphone sees as a plane wave in the far field, are shown in Fig. 16-2.

16.2 Pickup Patterns

Microphones are made with single- or multiple-pickup patterns and are named by the pickup pattern they employ. The pickup patterns and directional response characteristics of the various types of microphones are shown in Fig. 16-1.

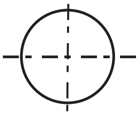
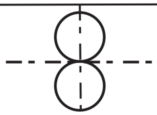
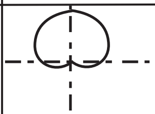
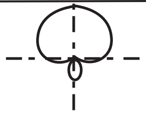
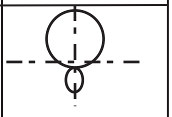
Microphone	Omnidirectional	Bidirectional	Directional	Supercardioid	Hypercardioid
Directional Response Characteristics					
Voltage output	$E = E_0$	$E = E_0 \cos\theta$	$E = \frac{E_0}{2}(1 + \cos\theta)$	$E = \frac{E_0}{2}[(\sqrt{3} - 1) + (3\sqrt{-3})\cos\theta]$	$E = \frac{E_0}{4}(1 + 3\cos\theta)$
Random energy Efficiency (%)	100	33	33	27	25
Front response / Back response	1	1	∞	3.8	2
Front random response / Total random response	0.5	0.5	0.67	0.93	0.87
Front random response / Back random response	1	1	7	14	7
Equivalent distance	1	1.7	1.7	1.9	2
Pickup angle (2 θ) for 3 dB attenuation	–	90°	130°	116°	100°
Pickup angle (2 θ) for 6 dB attenuation	–	120°	180°	156°	140°

Figure 16-1. Performance characteristics of various microphones.

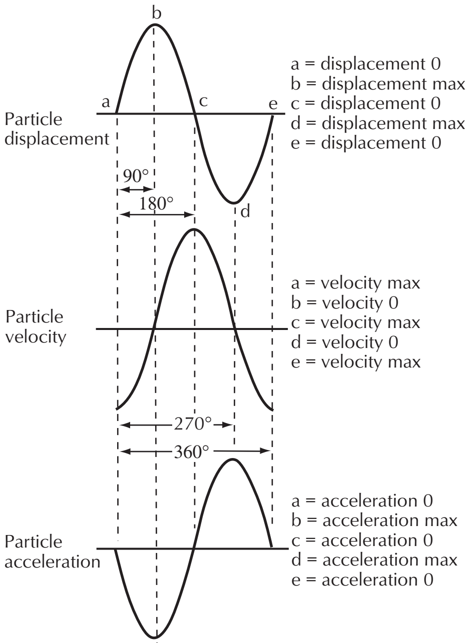

b
a
c
e
d
Particle displacement
a = displacement 0
b = displacement max
c = displacement 0
d = displacement max
e = displacement 0
90°
180°
Particle velocity
a = velocity max
b = velocity 0
c = velocity max
d = velocity 0
e = velocity max
270°
360°
Particle acceleration
a = acceleration 0
b = acceleration max
c = acceleration 0
d = acceleration max
e = acceleration 0

Figure 16-2. Air particle motion in a sound field, showing relationship to velocity and acceleration.

16.2.1 Omnidirectional Microphones

The omnidirectional, or spherical, polar response of the pressure microphones, Fig. 16-3 is created because the diaphragm is only exposed to the acoustic wave on the front side. Therefore, no cancellations are produced by having sound waves hitting both the front and rear of the diaphragm at the same time.

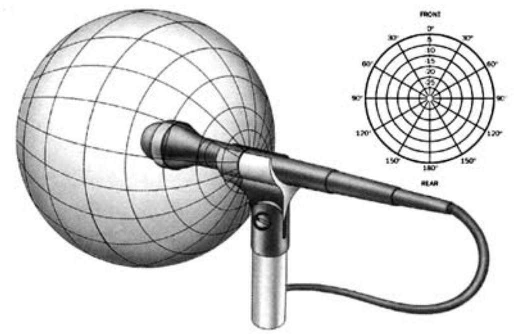


Figure 16-3. Omnidirectional pickup pattern. Courtesy Shure Incorporated.

Omnidirectional microphones become increasingly directional as the diameter of the microphone reaches the wavelength of the frequency in question, as shown in Fig. 16-4; therefore, the microphone should have the smallest diameter possible if omnidirectional characteristics are required at high frequencies. The characteristic that allows waves to bend around objects is known as *diffraction* and happens when the wavelength is long compared to the size of the object. As the wavelength approaches the size of the object, the wave cannot bend sharply enough and, therefore, passes by the object. The various responses start to diverge at the frequency at which the diameter of the diaphragm of the microphone, D, is approximately one-tenth the wavelength, λ, of the sound arriving as equation

$$D = \frac{\lambda}{10} \tag{16-1}$$

The frequency, f, at which the variation begins is

$$f = \frac{v}{10D} \tag{16-2}$$

where,

v is the velocity of sound in feet per second, or meters per second,

D is the diameter of the diaphragm in feet or meters.

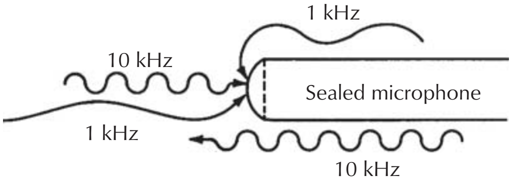

1 kHz
10 kHz
Sealed microphone
1 kHz
10 kHz

Figure 16-4. High-frequency directivity of an omnidirectional microphone.

For example, a ½ inch (1.27 cm) microphone will begin to vary from omnidirectional, though only slightly, at

$$\begin{aligned} f &= \frac{1130}{\left(10 \times \frac{0.5}{12}\right)} \\ &= 2712 \text{ Hz} \end{aligned}$$

and will be down approximately 3 dB at 10,000 Hz.

Omnidirectional microphones are capable of having a very flat, smooth frequency response over the entire audio spectrum because only the front of the diaphragm is exposed to the source, eliminating phase cancellations found in unidirectional microphones.

For smoothness of response the smaller they are, the better. The problem usually revolves around the smallest diaphragm possible versus the lowest signal-to-noise ratio, SNR, or put another way, the smaller the diaphragm, the lower the microphone sensitivity, therefore, the poorer the SNR.

Omnidirectional microphones have very little proximity effect. See Section 16.2.3.1 for a discussion on proximity effect.

Because the pickup pattern is spherical, the random energy efficiency is 100%, and the ratio of front response to back or side is 1:1, therefore signals from the sides or rear will have the same pickup sensitivity as from the front, giving a directivity index of 0 dB. This can be helpful in picking up wanted room characteristics or conversations around a table as when recording a symphony. However, it can be detrimental when in a noisy environment.

Omnidirectional microphones are relatively free from mechanical shock because the output at all frequencies is high; therefore, the diaphragm can be stiff. This allows the diaphragm to follow the magnet or stationary system it operates against when subjected to mechanical motion (see Section 16.3.3).

16.2.2 Bidirectional Microphones

A *bidirectional microphone* is one that picks up from the front and back equally well with little or no pickup from the sides. The field pattern, Fig. 16-5, is called a *figure eight*.

Because the microphone discriminates between the front, back, and sides, random energy efficiency is 33%. In other words, background noise, if it is in a reverberant field, will be 67% lower than with an omnidirectional microphone. The front-to-back response will still remain one; however, the front-to-side response will approach infinity, producing a directivity index of 4.8. This can be extremely useful when picking up two conversations on opposite sides of a table. Because of the increased directional capabilities of the microphone, pickup distance is 1.7 times greater before feedback in the direct field than for an omnidirectional microphone. The included pickup cone angle shown in Fig. 16-6 for 6 dB attenuation on a perfect bidirectional microphone is 120° off the front of the microphone and 120° off the rear of the microphone. Because of diffraction, this angle varies with frequency, becoming narrower as the frequency increases.

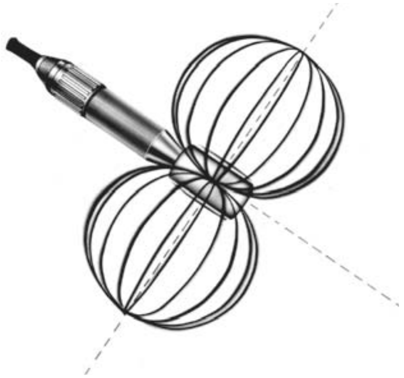


Figure 16-5. Bidirectional pickup pattern. Courtesy Sennheiser Electronic Corporation.

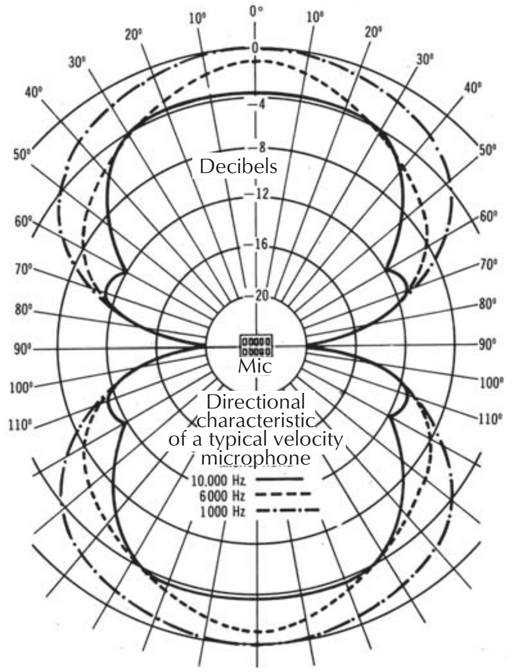


Figure 16-6. Polar pattern of a typical bidirectional ribbon velocity microphone showing the narrowing pattern at high frequencies.

16.2.3 Unidirectional Microphones

Unidirectional microphones have a greater sensitivity to sound pickup from the front than any other direction. The average unidirectional microphone has a front-to-back ratio of 20–30 dB; that is, it has 20–30 dB greater sensitivity to sound waves approaching from the front than from the rear.

Unidirectional microphones are usually listed as *cardioid* or *directional*, Fig. 16-7, *supercardioid*, Fig. 16-8, or *hypercardioid*, Fig. 16-9. The pickup pattern is called *cardioid* because it is heart shaped. Unidirectional microphones are the most commonly used microphones because they discriminate between signal and random unwanted noise. This has many advantages including:

- Less background noise,
- More gain before feedback especially when used in the direct field,
- Discrimination between sound sources.

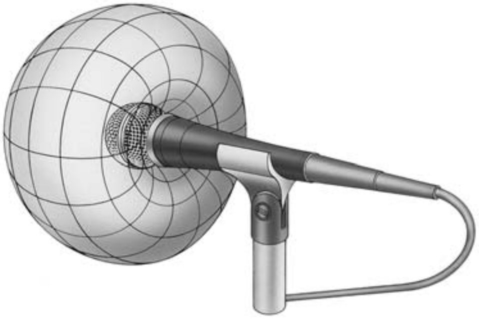


Figure 16-7. Cardioid pickup pattern. Courtesy Shure Incorporated.

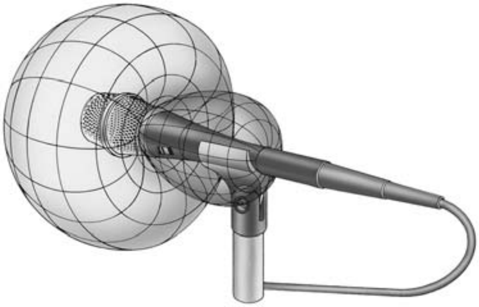


Figure 16-8. Supercardioid pickup pattern. Courtesy Shure Incorporated.

The cardioid pattern can be produced by one of two methods:

1. The first method combines the output of a pressure diaphragm and a pressure-gradient diaphragm, as shown in Fig. 16-10. Since the pressure-gradient diaphragm has a bidirectional pickup pattern and the pressure diaphragm has an omnidirectional pickup pattern, the wave hitting the front of the diaphragms add, while the wave hitting the rear of the diaphragm cancels as it is 180° out-of-phase with the rear pickup pattern of the pressure diaphragm. This method is expensive and seldom used for sound reinforcement or general-purpose microphones.

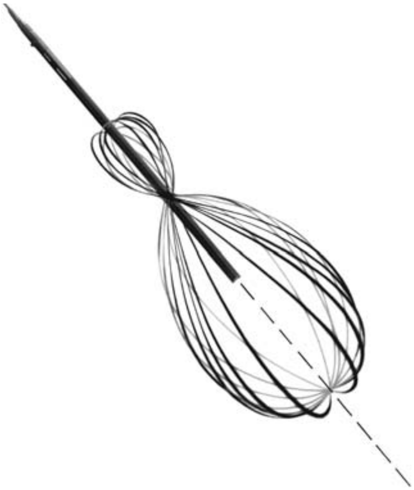


Figure 16-9. Hypercardioid pickup pattern. Courtesy Sennheiser Electronic Corporation.

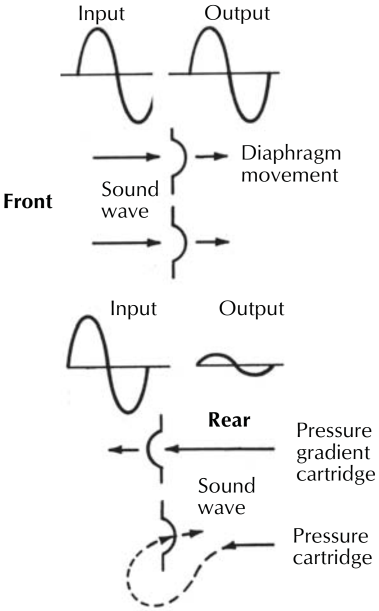
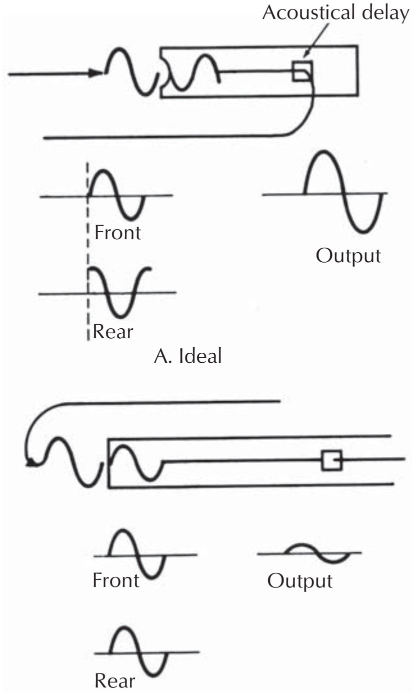


Figure 16-10. Two-diaphragm cardioid microphone.

2. The second and most widely used method of producing a cardioid pattern is to use a single diaphragm and acoustically delay the wave reaching the rear of the diaphragm. When a wave approaches from the front of the diaphragm, it first hits the front and then the rear of the diaphragm after traveling through an acoustical delay circuit, as shown in Fig. 16-11A. The pressure on the front of the diaphragm is at 0° while on the rear of the

diaphragm it is some angle between 0° and 180°, as shown in Fig. 16-11B. If the rear pressure was at 0°, the output would be 0. It would be ideal if the rear pressure were at 180° so that it could add to the input, doubling the output.



A. Ideal

B. Normal

Figure 16-11. Cardioid microphone employing acoustical delay.

The phase inversion is caused by the extra distance the wave has to travel to reach the back of the diaphragm. When the wave is coming from the rear of the microphone, it hits the front and back of the diaphragm at the same time and with the same polarity, therefore canceling the output.

The frequency response of cardioid microphones is usually rougher than an omnidirectional microphone due to the acoustical impedance path and its effects on the front wave response. The front and rear responses of a cardioid microphone are not the same. Although the front pattern may be essentially flat over the audio spectrum, the back response usually increases at low and high frequencies, as shown in Fig. 16-12.

Discrimination between the front and back response is between 15 and 30 dB in the mid frequencies and could be as little as 5–10 dB at the extreme ends, as shown in Fig. 16-12.

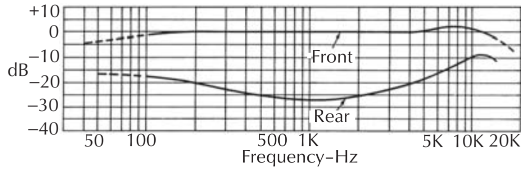


Figure 16-12. Frequency response of a typical cardioid microphone.

16.2.3.1 Proximity Effects

As the source is moved closer to the diaphragm, the low-frequency response increases due to the proximity effect, Fig. 16-13. The proximity effect[1] is created because at close source-to-microphone distance, the magnitude of the sound pressure on the front is appreciably greater than the sound pressure on the rear. In the vector diagram shown in Fig. 16-14A, the sound source was a distance greater than 2 ft from the microphone. The angle $2KD$ is found from D, which is the acoustic distance from the front to the rear of the diaphragm and $K = 2\pi/\lambda$. Fig. 16-14B shows the vector diagram when used less than 4 inches to the sound source.

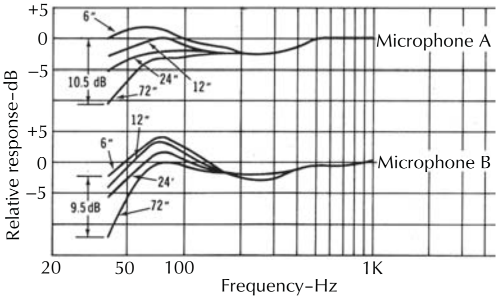
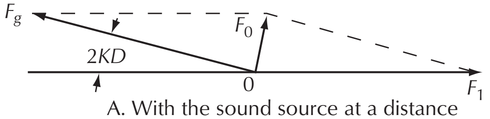
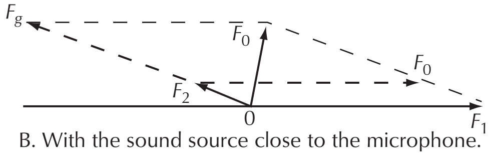


Figure 16-13. Proximity effect variations in response with distance between source and microphone for cardioid microphones. Courtesy Telex Electro-Voice.



A. With the sound source at a distance from the microphone.



B. With the sound source close to the microphone.

Figure 16-14. Vector diagram of a unidirectional microphone. Courtesy Telex Electro-Voice.

In both cases, force F_1, the sound pressure on the front of each diaphragm, is the same. Force F_2 is the force on the back of the diaphragm when the microphone is used at a distance from the sound source, and F_0 is the resultant force. The force F_2' on the back of the diaphragm is created by a close sound source. Laterally, the vector sum F_0' is considerably larger in magnitude than F_0 and therefore produces greater output from the microphone at low frequencies. This can be advantageous or disadvantageous. It is particularly useful when vocalists want to add low frequency to their voice or an instrumentalist to add low frequencies to the instrument. This is accomplished by varying the distance between the microphone and the sound source, increasing bass as the distance decreases.

16.2.3.1.1 Frequency Response

Frequency response is an important specification of unidirectional microphones and must be carefully analyzed and interpreted in terms of the way the microphone is to be used. If a judgment as to the sound quality of the microphone is made strictly from a single on-axis response, the influence of the proximity effect and off-axis response would probably be overlooked. A comparison of frequency response as a function of microphone-to-source distance will reveal that *all* unidirectional microphones experience a certain amount of proximity effect. In order to evaluate a microphone, this variation with distance is quite important.

When using a unidirectional microphone[2] in a handheld or stand-mounted configuration, it is conceivable that the performer will not always remain exactly on axis. Variations of ±45° often occur, and so a knowledge of the uniformity of response over such a range is important. The nature of these response variations is shown in Fig. 16-15. Response curves such as these give a better indication of this type of off-axis performance than polar response curves. The polar response curves are limited in that they are usually given for only a few frequencies, therefore, the complete spectrum is difficult to visualize.

For applications involving feedback control or noise rejection, the polar response or particular off-axis response curves, such as at 135° or 180°, are important. These curves can often be misleading due to the acoustic conditions and excitation signals used. Such measurements are usually made under anechoic conditions at various distances with sine-wave excitation. Looking solely at a rear response curve as a function of frequency is misleading since such a curve does not indicate the polar characteristic at any particular frequency, but only the level at one angle. Such curves also tend to give the impression of a rapidly fluctuating high-frequency discrimination. This sort of performance is to be expected since it is virtually impossible to design a microphone of practical size with a constant angle of best discrimination at high frequencies, Fig. 16-16. The principal factor influencing this variation in rear response is diffraction, which is caused by the physical presence of the microphone in the sound field. This diffraction effect is frequency dependent and tends to disrupt the ideal performance of the unidirectional phase-shift elements.

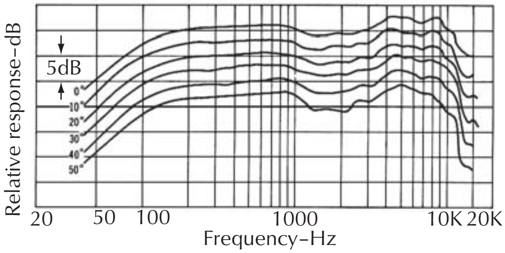


Figure 16-15. Variations in front response versus angular position. Note: Curves have been displaced by 2.5 dB for comparison purposes.

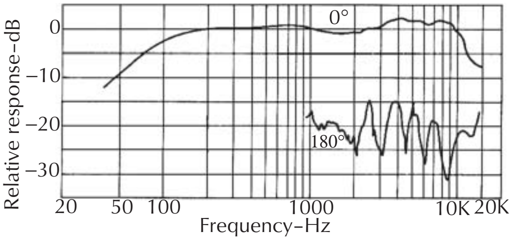


Figure 16-16. Typical fluctuations in high-frequency rear response for a cardioid microphone. Courtesy Shure Incorporated.

To properly represent this high-frequency off-axis performance, a polar response curve is of value, but it, too, can be confusing at high frequencies. The reason for this confusion can be seen in Fig. 16-17, where two polar response curves only 20 Hz apart are shown. The question that arises then is how can such performance be properly analyzed? A possible solution is to run polar response curves with bands of random noise such as ⅓ octaves of pink noise. Random noise is useful because of its averaging ability and because its amplitude distribution closely resembles program material.

Anechoic measurements are only meaningful as long as no large objects are in close proximity to the microphone. The presence of the human head in front of a

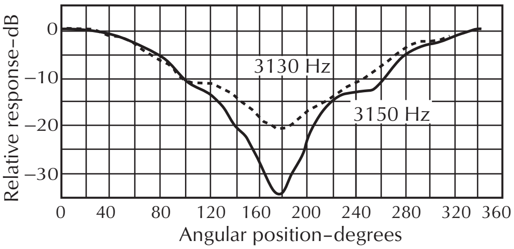


Figure 16-17. An example of rapid variations in high-frequency polar response for single-frequency excitation. Courtesy Shure Incorporated.

microphone will seriously degrade the effective high-frequency discrimination. An example of such degradation can be seen in Fig. 16-18 where a head object was placed 2 in (5 cm) in front of the microphone. (The two curves have not been normalized.) This sort of performance results from the head as a reflector and is a common cause of feedback as one approaches a microphone. This should not be considered as a shortcoming of the microphone, but rather as an unavoidable result of the sound field in which it is being used. At 180°, for example, the microphone sees, in addition to the source it is trying to reject, a reflection of that source some 2 in (5 cm) in front of its diaphragm. This phenomenon is greatly reduced at low frequencies because the head is no longer an appreciable obstacle to the sound field. It is thus clear that the effective discrimination of any unidirectional microphone is greatly influenced by the sound field in which it is used.

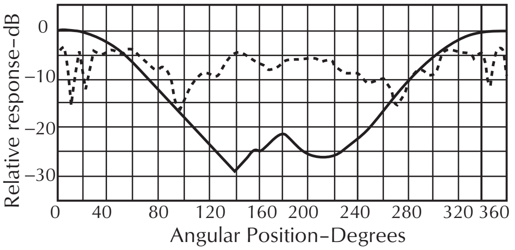


Figure 16-18. An example of a head obstacle on a polar response. Courtesy Shure Incorporated.

16.2.3.1.2 Types of Cardioid Microphones

Cardioid microphones are named by the way sound enters the rear cavity. The sound normally enters the rear of the microphone's cavity through single or multiple holes in the microphone housing, as shown in Fig. 16-19.

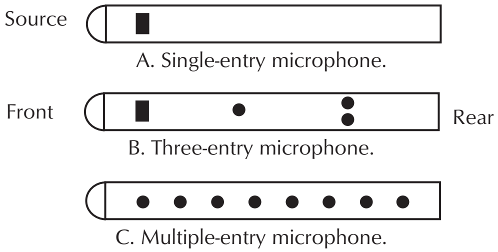

Source
A. Single-entry microphone.
Front
Rear
B. Three-entry microphone.
C. Multiple-entry microphone.

Figure 16-19. Three types of cardioid microphones.

16.2.3.1.3 Single-Entry Cardioid Microphones

All single-entrant cardioid microphones have the rear entrance port located at one distance from the rear of the diaphragm. The port location is usually within 1½ in (3.8 cm) of the diaphragm and can cause a large proximity effect. The Electro-Voice DS35 is an example of a single-entrant cardioid microphone, Fig. 16-20.

Figure 16-20. Electro-Voice DS35 single-entrant microphone. Courtesy Electro-Voice, Inc.

The low-frequency response of the DS35 varies as the distance from the sound source to the microphone decreases, Fig. 16-21. Maximum bass response is produced in close-up use with the microphone 1½ in (3.8 cm) from the sound source. Minimum bass response is experienced at distances greater than 2 ft (0.6 m). Useful effects can be created by imaginative application of the variable low-frequency response.

Another single-entrant microphone is the Shure SM-81.[3] The acoustical system of the microphone operates as a first-order gradient microphone with two sound openings, as shown in Fig. 16-22. Fig. 16-23 shows a simplified cross-sectional view of the transducer, and Fig. 16-23 indicates the corresponding electrical analog circuit of the transducer and preamplifier.

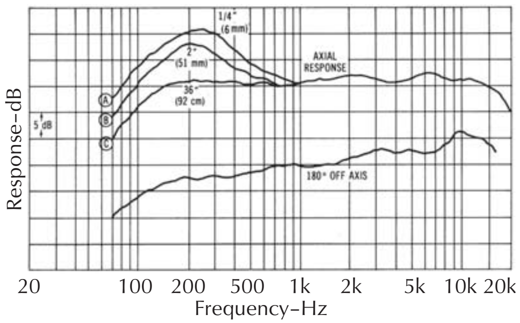


Figure 16-21. Frequency response versus distance for an Electro-Voice DS35 single-entrant cardioid microphone. Courtesy Electro-Voice, Inc.

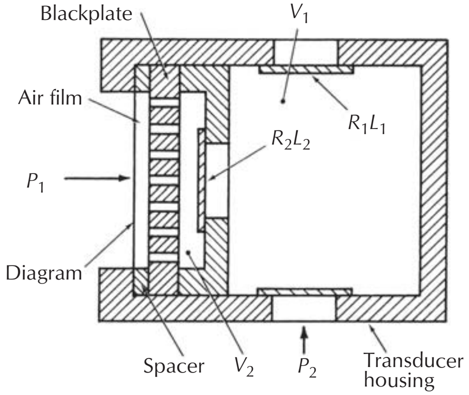


Figure 16-22. Simplified cross-sectional view of the Shure SM81 condenser transducer. Courtesy Shure Incorporated.

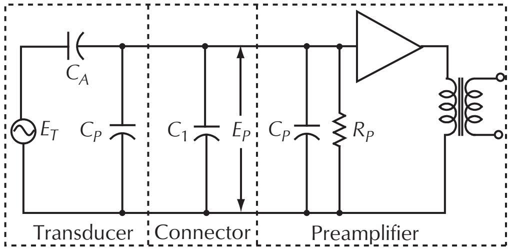


Figure 16-23. Electrical equivalent circuit of the Shure SM81 condenser transducer and preamplifier. Courtesy Shure Incorporated.

One sound opening, which is exposed to the sound pressure p_1, is represented by the front surface of the diaphragm. The other sound opening, or rear entry, consists of a number of windows in the side of the transducer housing where the sound pressure p_2 prevails. The diaphragm has an acoustical impedance Z_0, which also includes the impedance of the thin air film between the diaphragm and backplate. The sound pressure p_2 exerts its influence on the rear surface of the diaphragm via a screen mounted in the side windows of the transducer housing, having a resistance R_1 and inertance L_1, through the cavity V_1 with compliance C_1. A second screen has a resistance R_2 and inertance L_2, through a second cavity V_2 with compliance C_2, and finally through the perforations in the backplate.

The combination of circuit elements L_1, R_1, C_1, L_2, R_2, C_2 forms a ladder network with lossy inertances, and is called a *lossy ladder network*. The transfer characteristic of this network enforces a time delay on the pressure p_2 imparting directional (cardioid) characteristics for low and medium frequencies. At high frequencies the attenuation caused by the network is large, and the resulting pressure arriving at the back of the diaphragm due to p_2 is small. The microphone then operates much like an omnidirectional system under the predominant influence of p_1. At these frequencies directional characteristics are attained by diffraction of the sound around a suitably shaped transducer housing.

A rotary low-frequency response shaping switch allows the user to select between a flat and a 6 dB/octave roll-off at 100 Hz or an 18 dB/octave cutoff at 80 Hz. The 100 Hz roll-off compensates for the proximity effect associated with a 6 in (15 cm) source to microphone distance, while the 80 Hz cutoff significantly reduces most low-frequency disturbances with minimal effect on voice material. In the flat position the microphone has a 6 dB/octave electronic infrasonic roll-off, with –3 dB at 10 Hz to reduce the effects of inaudible low-frequency disturbances on microphone preamplifier inputs. Attenuation is provided for operation at high sound pressure levels (to 145 dB SPL) by means of a rotary capacitive switch (see Section 16.3.4.1).

A final example of a single entry cardioid microphone is the Shure Beta 57 supercardioid dynamic microphone, Fig.16-24. Both the Shure Beta 57 and the Beta 58 use neodymium magnets for hotter output and incorporate an improved shock mount.

16.2.3.1.4 Three-Entry Cardioid Microphones

The Sennheiser MD441 is an example of a three-entry cardioid microphone, Fig. 16-25. The low-frequency rear entry has a *d* (distance from center of diaphragm to the entry port) of about 2.8 inches (7 cm), the mid-frequency entry *d* is about 2.2 inches (5.6 cm) and the high-frequency entry *d* is about 1.5 inches (3.8 cm), with the transition in frequency occurring between 800 Hz and 1 kHz. Each entry consists of several holes around the microphone case rather than a single hole.

Figure 16-24. Shure Beta 57 dynamic microphone. Courtesy Shure Incorporated.

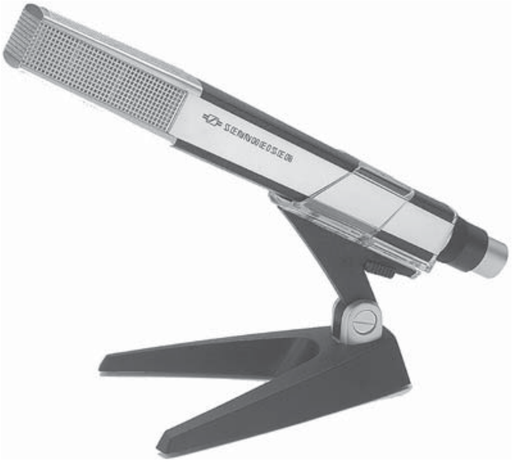


Figure 16-25. Sennheiser MD441 three-entry cardioid microphone. Courtesy Sennheiser Electronic Corporation.

This configuration is used for three reasons. By using a multiple arrangement of entry holes around the circumference of the microphone case into the low-frequency system, optimum front response and polar performance can be maintained, even though most of the entries may be accidentally covered when the microphone is handheld or stand mounted. The microphone has good proximity performance because the low-frequency entry ports are far from the diaphragm (4.75 inches) as well as the high-frequency entry having very little proximity influence at low frequencies. The two-entry configuration has a cardioid polar response pattern that provides a wide front working angle as well as excellent noise rejection and feedback control.

16.2.3.1.5 Multiple-Entry Cardioid Microphones

The Electro-Voice RE20 Continuously Variable-D microphone, Fig. 16-26, is an example of multiple-entry microphones. Multiple-entry microphones have many rear entrance ports. They can be constructed as single ports, all at a different distance from the diaphragm, or as a single continuous opening port. Each entrance is tuned to a different band of frequencies, the port closest to the diaphragm being tuned to the high frequencies, and the port farthest from the diaphragm being tuned to the low-frequency band. The greatest advantage of this arrangement is reduced-proximity effect because of the large distance between the source and the rear entry low-frequency port and mechanical crossovers are not as sharp and can be more precise for the frequencies in question.

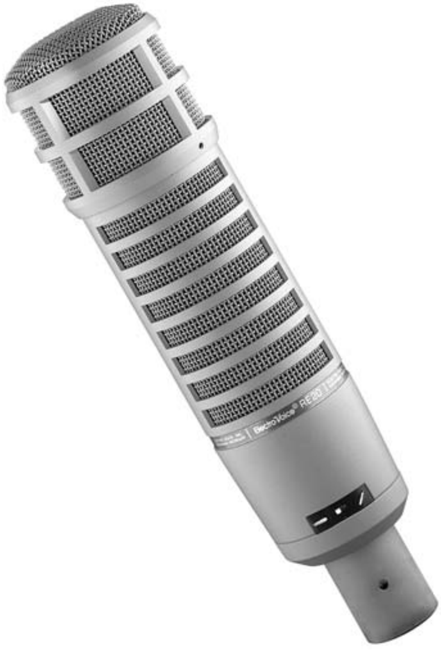


Figure 16-26. Electro-Voice RE20 multiple-entry (variable-D cardioid microphone). Courtesy Telex Electro-Voice.

As in many cardioid microphones, the RE20 has a low-frequency roll-off switch to reduce the proximity effect when close micing. Fig. 16-27 shows the wiring diagram of the RE20. By moving the red wire to either the 250 Ω or 50 Ω tap, the microphone output impedance can be changed. Note the "bass tilt" switch that, when open, reduces the series inductance and, therefore, the low-frequency response.

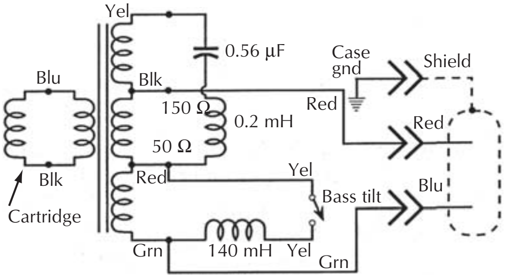

Yel
0.56 µF
Case gnd
Shield
Blu
Blk
150 Ω
0.2 mH
Red
Red
50 Ω
Yel
Blk
Red
Bass tilt
Blu
Cartridge
Grn
140 mH
Yel
Grn

Figure 16-27. Electro-Voice RE20 cardioid wiring diagram. Note "bass tilt" switch circuit and output impedance taps. Courtesy Telex Electro-Voice.

16.2.3.1.6 Two-Way Cardioid Microphones

In a two-way microphone system, the total response range is divided between a high-frequency and a low-frequency transducer, each of which is optimally adjusted to its specific range similar to a two-way loudspeaker system. The two systems are connected by means of a crossover network.

The AKG D-222EB schematically shown in Fig. 16-28 employs two coaxially mounted dynamic transducers. One is designed for high frequencies and is placed closest to the front grille and facing forward. The other is designed for low frequencies and is placed behind the first and facing rearward. The low-frequency transducer incorporates a hum-bucking winding to cancel the effects of stray magnetic fields. Both transducers are coupled to a 500 Hz inductive-capacitive-resistive crossover network that is electroacoustically phase corrected and factory preset for linear off-axis response. (This is essentially the same design technique used in a modern two-way loudspeaker system.)

The two-way microphone has a predominantly frequency-independent directional pattern, producing more linear frequency response at the sides of the microphone and far more constant discrimination at the rear of the microphone. Proximity effect at working distances down to 6 in (15 cm) is reduced because the distance between the microphone windscreen and the low frequency transducer is large.

The D-222EB incorporates a three-position bass-roll-off switch that provides 6 dB or 12 dB attenuation at 50 Hz. This feature is especially useful in speech applications and in acoustically unfavorable environments with excessive low-frequency ambient noise, reverberation, or feedback.

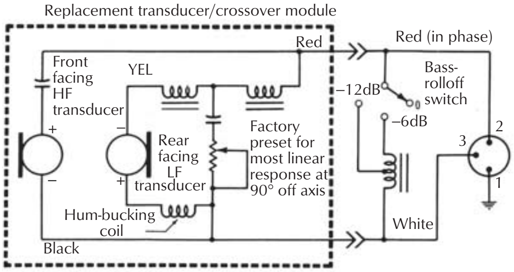

Replacement transducer/crossover module
Red
Red (in phase)
Front facing HF transducer
YEL
−12dB
Bass-rolloff switch
−6dB
Factory preset for most linear response at 90° off axis
Rear facing LF transducer
2
3
1
Hum-bucking coil
White
Black

Figure 16-28. Schematic of an AKG D222EB two-way cardioid microphone. Courtesy AKG Acoustics, Inc.

16.3 Types of Transducers

16.3.1 Carbon Microphones

One of the earliest types of microphones, the *carbon microphone*, is still found in old telephone handsets. It has very limited frequency response, is very noisy, has high distortion, and requires a hefty dc power supply. A carbon microphone[4] is shown in Fig. 16-29 and operates in the following manner.

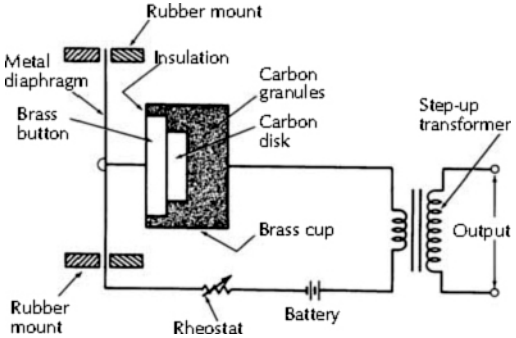

Rubber mount
Metal diaphragm
Insulation
Carbon granules
Brass button
Carbon disk
Step-up transformer
Brass cup
Output
Rubber mount
Rheostat
Battery

Figure 16-29. Connection and construction of a single-button carbon microphone.

Several hundred small carbon granules are held in close contact in a brass cup called a *button* that is attached to the center of a metallic diaphragm. Sound waves striking the surface of the diaphragm disturb the carbon granules, changing the contact resistance between their surfaces. A battery or dc power source is connected in series with the carbon button and the primary of an audio impedance-matching transformer.

The change in contact resistance causes the current from the power supply to vary in amplitude, resulting in a current waveform similar to the acoustic waveform striking the diaphragm.

The impedance of the carbon button is low so a step-up transformer is used to increase the impedance and voltage output of the microphone and to eliminate dc from the output circuit.

16.3.2 Crystal and Ceramic Microphones

Crystal and *ceramic* microphones were once popular because they were inexpensive and their high-impedance high-level output allowed them to be connected directly to the input grid of a tube amplifier. They were most popular in use with home tape recorders where microphone cables were short and input impedances high.

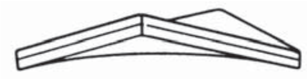
Crystal and ceramic microphones[5] operate as follows: piezoelectricity is "pressure electricity" and is a property of certain crystals such as Rochelle salt, tourmaline, barium titanate, and quartz. When pressure is applied to these crystals, electricity is generated. Present-day commercial materials such as ammonium dihydrogen phosphate (ADP), lithium sulfate (LN), dipotassium tartrate (DKT), potassium dihydrogen phosphate (KDP), lead zirconate, and lead titanate (PZT) have been developed for their piezoelectric qualities. Ceramics do not have piezoelectric characteristics in their original state, but the characteristics are introduced in the materials by a polarizing process. In piezoelectric ceramic materials the direction of the electrical and mechanical axes depends on the direction of the original dc polarizing potential. During polarization a ceramic element experiences a permanent increase in dimensions between the poling electrodes and a permanent decrease in dimension parallel to the electrodes.

The crystal element can be cut as a bender element that is only affected by a bending motion or as a twister element that is only affected by a twisting motion, Fig. 16-30.

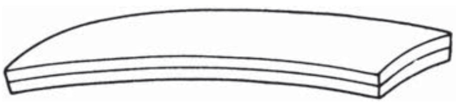
The internal capacitance of a crystal microphone is about 0.03 µF for the diaphragm-actuated type and 0.0005–0.015 µF for the sound-cell type.

The *ceramic* microphone operates like a crystal microphone except it employs a barium titanate slab in the form of a ceramic, giving it better temperature and humidity characteristics.

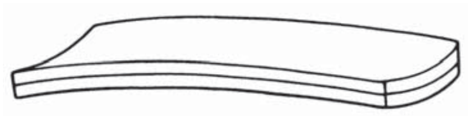
Crystal and ceramic microphones normally have a frequency response from 80 to 6500 Hz but can be made to have a flat response to 16 kHz. Their output impedance is about 100 kΩ, and they require a minimum load of 1–5 MΩ to produce a level of about –30 dB re 1 V/Pa.



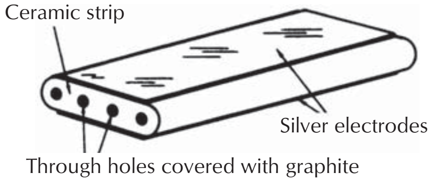
A. Crystal twister bimorph.



B. Ceramic bender bimorph.



C. Crystal bender bimorph.



D. Multimorph.

Figure 16-30. Curvatures of bimorphs and multimorph. Courtesy Clevite Corp., Piezoelectric Division.

16.3.3 Dynamic Microphones

The *dynamic microphone* is also referred to as a *pressure* or *moving-coil* microphone. It employs a small diaphragm and a voice coil, moving in a permanent magnetic field. Sound waves striking the surface of the diaphragm cause the coil to move in the magnetic field, generating a voltage proportional to the sound pressure at the surface of the diaphragm.

In a dynamic pressure unit, Fig. 16-31, the magnet and its associated parts (magnetic return, pole piece, and pole plate) produce a concentrated magnetic flux of approximately 10,000 G in the small gap.

The diaphragm, a key item in the performance of a microphone, supports the voice coil centrally in the magnetic gap, with only 0.006 inch clearance.

An omnidirectional diaphragm and voice-coil assembly is shown in Fig. 16-32. The compliance section has two hinge points with the section between them made up of tangential corrugated triangular sections that stiffen this portion and allow the diaphragm to move in and out with a slight rotating motion. The hinge points are designed to permit high-compliance action. A spacer supports the moving part of the diaphragm away from the top pole plate to

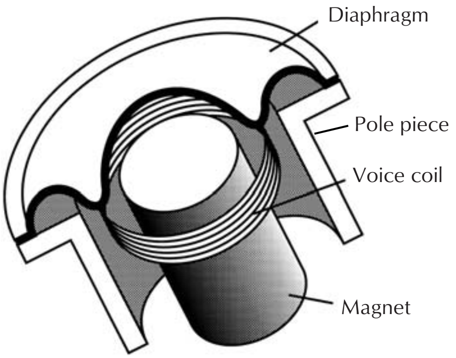


Figure 16-31. A simplified drawing of a dynamic microphone. Courtesy Shure Incorporated.

provide room for its movement. The cementing flat is bonded to the face plate. A stiff hemispherical dome is designed to provide adequate acoustical capacitance. The coil seat is a small step where the voice coil is mounted, centered, and bonded on the diaphragm.

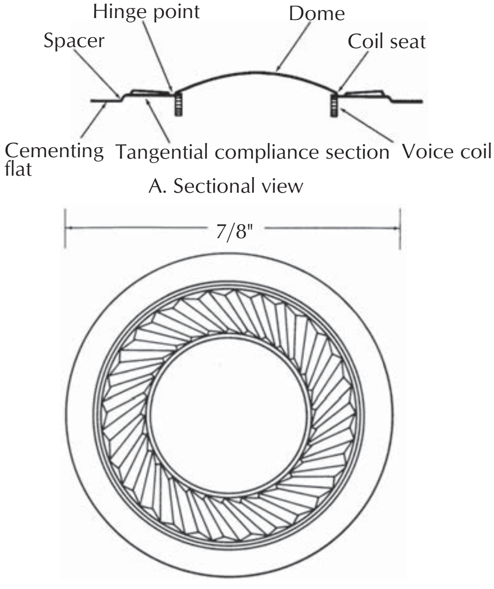


Figure 16-32. Omnidirectional diaphragm and voice coil assembly.

Early microphones had aluminum diaphragms that were less than 1 mil (0.001 in) thick. Aluminum is light-weight, easy to form, maintains its dimensional stability, and is unaffected by extremes in temperature or humidity. Unfortunately, being only 1 mil thick makes the diaphragms fragile. When it is touched or otherwise deformed by excessive pressure, an aluminum diaphragm is dead.

Mylar™, a polyester film manufactured by the DuPont Company, is commonly used for diaphragms. Mylar is a unique plastic. Extremely tough, it has high tensile strength, high resistance to wear, and outstanding flex life. Mylar™ diaphragms have been cycle tested with temperature variations from −40°F to +170°F (−40°C to +77°C) over long periods without any impairment to the diaphragm. Since Mylar™ is extremely stable, its properties do not change within the temperature and humidity range in which microphones are used.

The specific gravity of Mylar™ is approximately 1.3 as compared to 2.7 for aluminum so a Mylar™ diaphragm may be made considerably thicker without upsetting the relationship of the diaphragm mass to the voice-coil mass.

Mylar™ diaphragms are formed under high temperature and high pressure, a process in which the molecular structure is formed permanently to establish a *dimensional memory* that is highly retentive. Unlike aluminum, Mylar™ diaphragms will retain their shape and dimensional stability although they may be subjected to drastic momentary deformations.

The voice coil weighs more than the diaphragm so it is the controlling part of the mass in the diaphragm voice-coil assembly. The voice coil and diaphragm mass (analogous to inductance in an electrical circuit) and compliance (analogous to capacitance), make the assembly resonate at a given frequency as any tuned electrical circuit. The free-cone resonance of a typical undamped unit is in the region of 350 Hz.

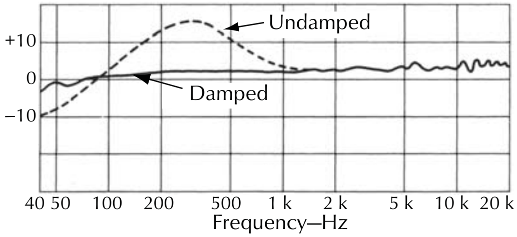


Figure 16-33. Diaphragm and voice-coil assembly response curve.

If the voice coil were left undamped, the response of the assembly would peak at 350 Hz, Fig. 16-33. The resonant characteristic is damped out by the use of an acoustic resistor, a felt ring that covers the openings in the centering ring behind the diaphragm. This is analogous to electrical resistance in a tuned circuit. While this reduces the peak at 350 Hz, it does not fix the droop below 200 Hz. Additional acoustical resonant devices are used inside the microphone case to correct the drooping. A cavity behind the unit (analogous to capacitance) helps resonate at the low frequencies with the mass (inductance) of the diaphragm and voice-coil assembly.

Another tuned resonant circuit is added to extend the response down to 35 Hz. This circuit, tuned to about 50 Hz, is often a tube that couples the inside cavity of the microphone housing to the outside, Fig. 16-34.

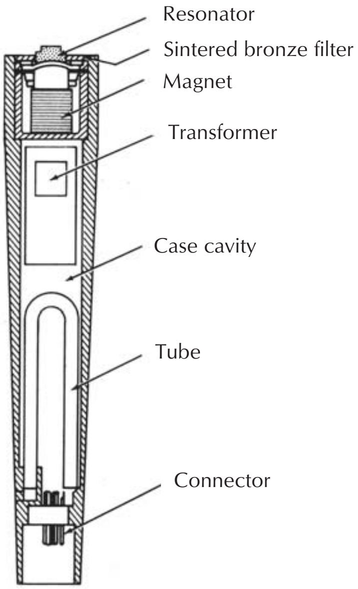


Figure 16-34. Omnidirectional microphone cross-section view.

The curvature of the diaphragm dome provides stiffness, and the air cavity between it and the dome of the pole piece form an acoustic capacitance. This capacitance resonates with the mass (inductance) of the assembly to extend the response up to 20 kHz.

To control the high-frequency resonance, a nonmagnetic filter is placed in front of the diaphragm, creating an acoustic resistance, Fig. 16-34. The filter is also an effective mechanical protection device. The filter prevents dirt particles, magnetic chips, and moisture from gravitating to the inside of the unit. Magnetic chips, if allowed to reach the magnetic gap area, eventually will accumulate on top of the diaphragm and impair the frequency response. It is possible for such chips to pin the diaphragm to the pole piece, rendering the microphone inoperative.

Fig. 16-35 illustrates the effect of a varying sound pressure on a moving-coil microphone. For this simplified explanation, assume that a massless diaphragm voice-coil assembly is used. The acoustic waveform, Fig. 16-35A, is one cycle of an acoustic waveform, where *a* indicates atmospheric pressure *AT*; and *b* represents atmospheric pressure plus a slight overpressure increment Δ or $AT + \Delta$.

The electrical waveform output from the moving-coil microphone, Fig. 16-35B, does not follow the phase of the acoustic waveform because at maximum pressure, $AT + \Delta$ or *b*, the diaphragm is at rest (no velocity). Further, the diaphragm and its attached coil reach maximum velocity, hence maximum electrical amplitude—at point *c* on the acoustic waveform. This is of no consequence unless another microphone is being used along with the moving-coil microphone where the other microphone does not see the same 90° displacement. Due to this phase displacement, condenser microphones should not be mixed with moving-coil or ribbon microphones when micing the same source at the same distance. (Sound pressure can be proportional to velocity in many practical cases.)[6]

A steady overpressure which can be considered an acoustic square wave, Fig. 16-35C, would result in the output shown in Fig. 16-35D. As the acoustic pressure rises from *a* to *b*, it has velocity, Fig. 16-35, creating a voltage output from the microphone. Once the diaphragm reaches its maximum displacement at *b*, and stays there during the time interval represented by the distance between *b* and *c*, voice-coil velocity is zero so electrical output voltage ceases and the output returns to zero. The same situation repeats itself from *c* to *e* and from *e* to *f* on the acoustic waveform. As can be seen, a moving-coil microphone cannot reproduce a square wave.

Another interesting theoretical consideration of the moving-coil microphone mechanism is shown in Fig. 16-36. Assume a sudden transient condition. Starting at *a* on the acoustic waveform, the normal atmospheric pressure is suddenly increased by the first wavefront of a new signal and proceeds to the first overpressure peak, $AT + \Delta$ or *b*. The diaphragm will reach a maximum velocity halfway to *b* and then return to zero velocity at *b*. This will result in a peak, *a'*, in the electrical waveform. From *b* on, the acoustic waveform and the electrical waveform will proceed as before, cycle for cycle, but 90° apart.

In this special case, peak *a'* does not follow the input precisely so it is something extra. It will probably be swamped out by other problems (especially mass) encountered in a practical moving-coil microphone. It

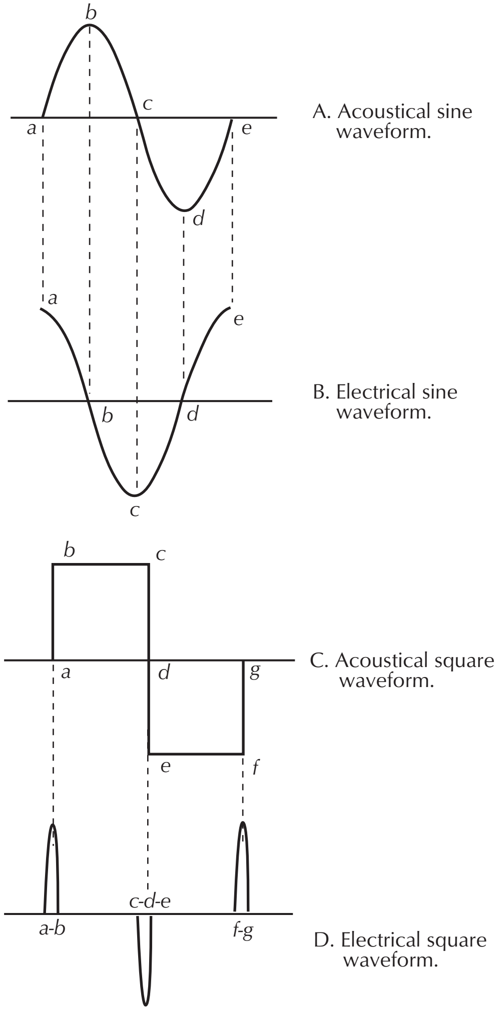

b
c
a
e
A. Acoustical sine waveform.
d
a
e
B. Electrical sine waveform.
b
d
c
b
c
C. Acoustical square waveform.
a
d
g
e
f
c-d-e
a-b
f-g
D. Electrical square waveform.

Figure 16-35. Effect of a varying sound pressure on a moving-coil microphone.

does illustrate that even with a "perfect," massless, moving-coil microphone, "perfect" electrical waveforms will not be produced.

When sound waves vibrate the diaphragm, the voice coil has a voltage induced in it proportional to the magnitude and at the frequency of the vibrations. The voice coil and diaphragm have some finite mass and any mass has inertia that causes it to want to stay in the condition it is in—namely, in motion or at rest. If the stationary part of the diaphragm-magnet structure is moved in space, the inertia of the diaphragm and coil causes them to try to remain fixed in space. Therefore, there will be relative motion between the two parts with a resultant electrical output. An electrical output can be obtained in two ways, by motion of the diaphragm from airborne acoustical energy or by motion of the magnet circuit by structure-borne vibration. The diaphragm motion is the desired output, while the structure-borne vibration is undesired.

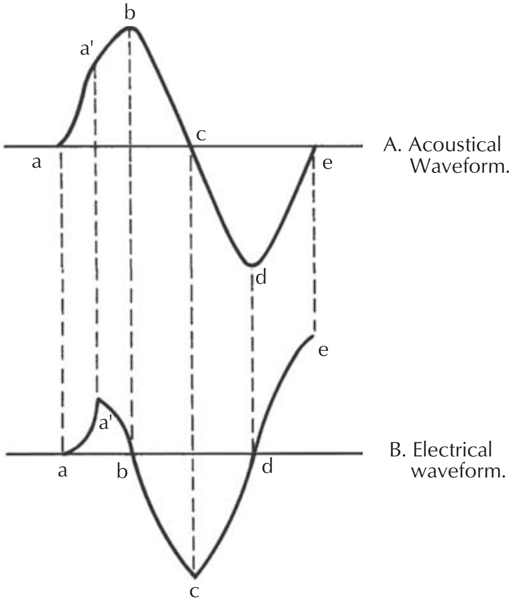


Figure 16-36. Effect of a transient condition on a moving-coil microphone.

Several things may be tried to eliminate the undesired output. The mass of the diaphragm and voice coil may be reduced, but there are practical limits, or the frequency response may be limited mechanically with stiffer diaphragms or electronically with filter circuits. However limited response makes the microphone unsuitable for broad-range applications.

16.3.3.1 Unidirectional Microphones

To reject unwanted acoustical noise such as signals emanating from the sides or rear of the microphone, unidirectional microphones are used, Fig. 16-7. Unidirectional microphones are much more sensitive to vibration relative to their acoustic sensitivity than omnidirectional types. Fig. 16-37 shows a plot of vibration sensitivity versus frequency for a typical omnidirectional and unidirectional microphone with the levels normalized with respect to acoustical sensitivity.

The vibration sensitivity of the unidirectional microphone is about 15 dB higher than the omnidirectional and has a peak at about 150 Hz. The peak gives a clue to help explain the difference.

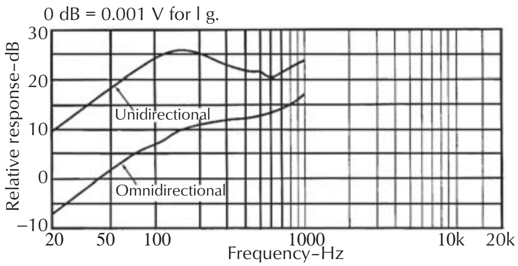


Figure 16-37. Vibration sensitivity of microphone cartridge.

Unidirectional microphones are usually differential microphones; that is, the diaphragm responds to a pressure differential between its front and back surfaces. The oncoming sound wave is not only allowed to reach the front of a diaphragm but, through one or more openings and appropriate acoustical phase-shift networks, reaches the rear of the diaphragm. At low frequencies, the net instantaneous pressure differential causing the diaphragm to move is small compared to the absolute sound pressure, Fig. 16-38. Curve A is the pressure wave that arrives at the front of the diaphragm. Curve B is the pressure wave that reaches the rear of the diaphragm after a slight delay due to the greater distance the sound had to travel to reach the rear entry and some additional phase shift it encounters after entering. The net pressure actuating the diaphragm is curve C, which is the instantaneous difference between the two upper curves. In a typical unidirectional microphone, the differential pressure at 100 Hz will be about one-tenth of the absolute pressure or 20 dB down from the pressure an omnidirectional microphone would experience.

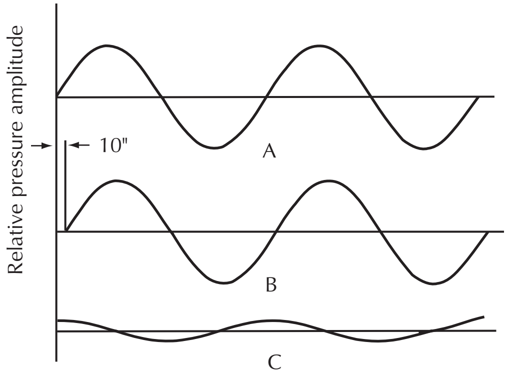


Figure 16-38. Differential pressure at low frequencies on unidirectional microphones.

To obtain good low-frequency response, a reasonable low-frequency electrical output is required from a unidirectional microphone. To accomplish this, the diaphragm must move more easily for a given sound pressure. Some of this is accomplished by reducing the damping resistance to less than one-tenth used in an omnidirectional microphone. This reduction in damping increases the motion of the mechanical resonant frequency of the diaphragm and voice coil, around 150 Hz in Fig. 16-37, making the microphone much more acceptable to structure-borne vibrations. Since the diaphragm of an omnidirectional microphone is much more heavily damped, it will respond less to inertial or mechanical vibration forces.

To eliminate unwanted external low-frequency noise from effecting a unidirectional microphone, some kind of isolation such as a microphone shock mount is required to prevent the microphone cartridge from experiencing mechanical shock and vibration.

16.3.4 Capacitor Microphones

In a *capacitor* or *condenser* microphone the sound pressure level varies the head capacitance of the microphone by deflecting one or two plates of the capacitor, causing an electrical signal that varies with the acoustical signal. The varying capacitance can be used to modulate an RF signal that is later demodulated or can be used as one leg of a voltage divider, Fig. 16-39, where *R* and *C* form the voltage divider of the power supply ++ to –.

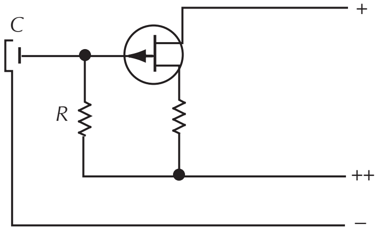


Figure 16-39. Voltage divider type of capacitor microphone.

The head of most capacitor microphones consists of a small two-plate 40–50 pF capacitor. One of the two plates is a stretched diaphragm; the other is a heavy back plate or center terminal, Fig. 16-39. The back plate is insulated from the diaphragm and spaced approximately 1 mil (0.001 in) from, and parallel to, the rear surface of the diaphragm. Mathematically the output from the head may be calculated as

$$E_O = \frac{E_p a^2 P}{8dt} \qquad (16\text{-}3)$$

where,

E_O is the output in volts,

E_p is the dc polarizing voltage in volts,

a is the radius of active area of the diaphragm in centimeters,

P is the pressure in dynes per square centimeter,

d is the spacing between the back plate and diaphragm in centimeters,

t is the diaphragm tension in dynes per centimeter.

Many capacitor microphones operate with an equivalent noise level of 15–30 dB SPL. Although a 20–30 dB SPL is in the range of a well-constructed studio, a 20–30 dB microphone equivalent noise is not masked by room noise because room noise occurs primarily at low frequencies and microphone noise at high frequency as hiss.

In the past the quality of sound recordings was limited by the characteristics of the analog tape and record material, apart from losses induced by the copying and pressing procedures. Tape saturation, for instance, created additional harmonic and disharmonic distortion components, which affected the recording fidelity at high levels, whereas the linearity at low and medium levels was quite acceptable. The onset of these distortions was rather soft and extended to a wide level range that makes it difficult to determine the threshold of audibility.

The distortion characteristics of the standard studio condenser microphones is adequate for operation with analog recording equipment. Although exhibiting a high degree of technical sophistication, these microphones show individual variations in the resolution of complex tonal structures, due to their specific frequency responses and directivity patterns and nonlinear effects inherent to these microphones.

These properties were mostly concealed by the distortions superimposed by the analog recording and playback processing. But the situation has changed essentially since the introduction of digital audio. The conversion of analog signals into digital information and vice versa is carried out very precisely, especially at high signal levels. Due to the linear quantization process the inherent distortions of digital recordings virtually decrease at increasing recording levels, which turns former distortion behavior upside down. This new reality, which is in total contrast to former experience with analog recording technique, contributes mostly to the fact that the specific distortion characteristics of the microphone may become obvious, whereas they have been masked previously by the more significant distortions of analog recording technique.

Another feature of digital audio is the enlarged dynamic range and reduced noise floor. Unfortunately, due to this improvement, the inherent noise of the microphones may become audible, because it is no longer covered up by the noise of the recording medium.

The capacitor microphone has a much faster rise time than the dynamic microphone because of the significantly lower mass of the moving parts (diaphragm versus diaphragm/coil assembly). The capacitor rise time rises from 10% of its rise time to 90% in approximately 15 μs, while the rise time for the dynamic microphone is in the order of 40 μs.

Capacitor microphones generate an output electrical waveform in step or phase with the acoustical waveform and can be adapted to measure essentially dc overpressures, Fig. 16-40.

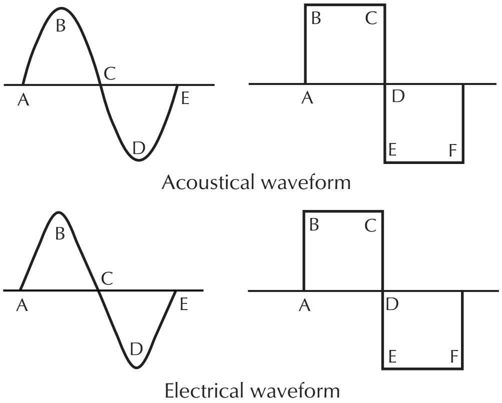


Figure 16-40. Capacitor microphone acoustic wave and electrical signals. Note the in-phase condition.

Some advantages of capacitor microphones are:

- Small, low-mass rigid diaphragms that reduce vibration pickup.
- Smooth, extended-range frequency response.
- Rugged—capable of measuring very high sound pressure levels (rocket launches).
- Low noise (which is partially cancelled by the need for electronics).
- Small head size, which provides low diffraction interference.

16.3.4.1 Voltage Divider Capacitor Microphone

Voltage divider-type capacitor microphones require a preamplifier as an integral part of the housing and a source of polarizing voltage for the head, and a source of power.

A high-quality capacitor microphone, the Sennheiser K6 Modular Condenser Microphone Series is suitable for recording studios, television and radio broadcast, motion picture studios, and stage and concert hall applications, as well as high-quality commercial sound installations.

The K6/ME62 series is a *capacitor microphone system,* Fig. 16-41, that uses AF circuitry with field-effect transistors so it has a low noise level (15 dB per DIN IEC 651), high reliability, and lifelong stability. Low current consumption at low voltage and phantom circuit powering permit feeding the microphone supply voltage via a standard two-conductor shielded audio cable or an internal AA battery.

Figure 16-41. Modular microphone system with an omnidirectional cartridge utilizing a voltage divider circuit. Courtesy Sennheiser Electronic Corporation.

The K6 offers interchangeable capsules, allowing the selection of different response characteristics from omnidirectional to cardioid to shotgun to adapt the microphone to various types of environments and recording applications.

Because of the new PCM recorders, signal-to-noise ratio (SNR) has reached a level of 90 dB, requiring capacitor microphones to increase their SNR level to match the recorder. The shotgun K6 series microphone, Fig. 16-42, has an equivalent noise level of 16 dB (DIN

Figure 16-42. The same microphone shown in Fig. 16-41 but with a shotgun cartridge. Courtesy Sennheiser Electronic Corporation.

As in most circuitry, the input stage of a voltage divider-type capacitor microphone has the most effect on noise, Fig. 16-43. It is important that the voltage on the transducer does not change. This is normally accomplished by controlling the input current. In the circuit of Fig. 16-43, the voltages V_{in}, V_o, and V_D are within 0.1% of each other. Noise, which might come into the circuit as V_{in} through the operational amplifier, is only $1/91$ of the voltage V_o.

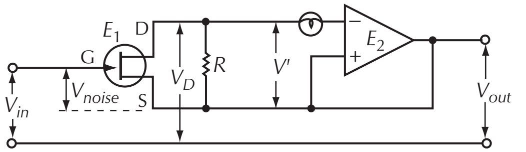


Figure 16-43. Simplified schematic of an AKG C-460B microphone input circuit. Courtesy AKG Acoustics.

Preattenuation, that is, attenuation between the capacitor and the amplifier, can be achieved by connecting parallel capacitors to the input, by reducing the input stage gain by means of capacitors in the nega-

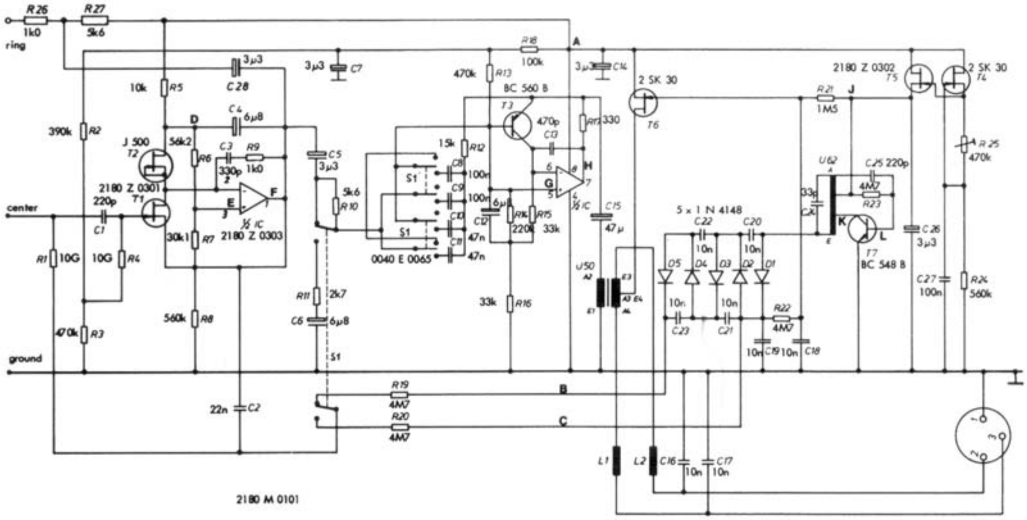

R26
1k0
ring
R27
5k6
center
ground
R1
10G
R4
10G
C1
220p
R2
390k
R3
470k
R5
10k
J 500
T2
2180 Z 0301
T1
D
56k2
R6
C28
3µ3
C4
6µ8
C3
330p
R9
1k0
E
F
½ IC
2180 Z 0303
30k1
R7
560k
R8
22n
C2
3µ3
C7
C5
3µ3
5k6
R10
R11
2k7
C6
6µ8
S1
R19
4M7
R20
4M7
B
C
15k
R12
C8
100n
C9
100n
C10
47n
C11
47n
0040 E 0065
470k
R13
R18
100k
A
BC 560 B
T3
470p
C13
C12
6µ8
R14
220k
R15
33k
R16
33k
G
H
½ IC
R17
330
3µ3
C14
C15
47µ
U50
2 SK 30
T6
5 x 1 N 4148
C22
10n
C20
10n
D5
D4
D3
D2
D1
C23
10n
C21
10n
R22
4M7
C19
10n
C18
10n
L1
L2
C16
10n
C17
10n
R21
1M5
J
U62
C24
33p
C25 220p
R23
4M7
K
L
T7
BC 548 B
2180 Z 0302
T5
2 SK 30
T4
R25
470k
C26
3µ3
C27
100n
R24
560k
1
2
3
2180 M 0101

Figure 16-44. Schematic of an AKG C-460B microphone. Courtesy AKG Acoustics.

tive feedback circuit, or by reducing the polarizing voltage to one-third its normal value and by using a resistive voltage divider in the audio line. Fig. 16-44 is the schematic for the AKG C-460B microphone.

16.3.4.2 Phantom Power for Capacitor Microphones

A common way to supply power for capacitor microphones is with a *phantom circuit*. *Phantom* or *simplex powering* is supplying power to the microphone from the input of the following device such as a preamplifier, mixer, or console.

Most capacitor microphone preamplifiers can operate on any voltage between 9 Vdc and 52 Vdc because they incorporate an internal voltage regulator. The preamplifier supplies the proper polarizing voltage for the capacitor capsule plus impedance matches the capsule to the balanced low-impedance output.

Standard low-impedance, balanced microphone input receptacles are easily modified to simplex both operating voltage and audio output signal, offering the following advantages in reduced cost and ease of capacitor microphone operation:

- Special external power supplies and separate multiconductor cables formerly required with capacitor microphones can be eliminated.
- The B+ supply in associated recorders, audio consoles, and commercial sound amplifiers can be used to power the microphone directly.
- Dynamic, ribbon, and capacitor microphones can be used interchangeably on standard, low-impedance, balanced microphone circuits.
- Dynamic, ribbon, and self-powered capacitor microphones may be connected to the modified amplifier input without defeating the microphone operating voltage.
- Any recording, broadcast, and commercial installation can be inexpensively upgraded to capacitor microphone operation using existing, two-conductor microphone cables and electronics.

Phantom circuit use requires only that the microphone operating voltage be applied equally to pins 2 and 3 of the amplifier low-impedance (normally an XLR input) receptacle. Pin 1 remains ground and circuit voltage minus. The polarity of standard microphone cable wiring is not important except for the usual audio polarity requirement (see Section 16.5.3). Two equally effective methods of amplifier powering can be used:

1. Connect an amplifier B+ supply of 9–12 V directly to the ungrounded center tap of the microphone input transformer, as shown in Fig. 16-45. A series-dropping resistor is required for voltages between 12 and 52 V. Fig. 16-46 is a typical resistor value chart. A chart can be made for any microphone if the current is known for a particular voltage.
2. A two-resistor, artificial center powering circuit is required when the microphone input transformer is not center-tapped, or when input attenuation

networks are used across the input transformer primary. Connect a B+ supply of 9–12 V directly to the artificial center of two 332 Ω, 1% tolerance precision resistors, as shown in Fig. 16-47. Any transformer center tap should not be grounded. For voltages between 12 and 52 V, double the chart resistor value of Fig. 16-46.

Any number of capacitor microphones may be powered by either method from a single B+ source according to the current available. Use the largest resistor value shown (R_v *max*) for various voltages in Fig. 16-46 for minimum current consumption.

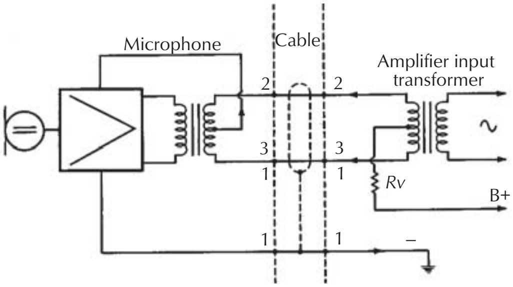


Figure 16-45. Direct center-tap connection method of phantom powering capacitor microphones. Courtesy AKG Acoustics.

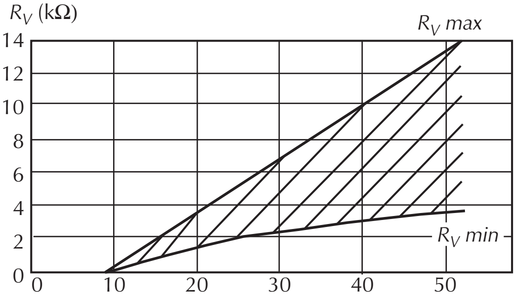


Figure 16-46. Dropping resistor value chart for phantom powering AKG C-451E microphones. Courtesy AKG Acoustics.

16.3.4.3 Capacitor Radio-Frequency, Frequency-Modulated Microphones

A *frequency-modulated microphone* is a capacitor microphone that is connected to a radio-frequency (RF) oscillator. Pressure waves striking the diaphragm cause variations in the capacity of the microphone head that frequency modulates the oscillator. The output of the modulated oscillator is passed to a discriminator and amplified in the usual manner.

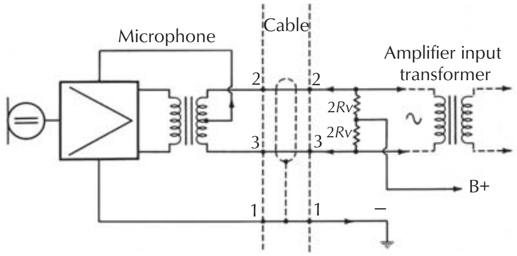


Figure 16-47. Artificial center tap connected method of powering capacitor microphones. Courtesy AKG Acoustics.

Capacitor microphones using an RF oscillator are not entirely new to the recording profession, but since the advent of solid-state devices, considerable improvement has been achieved in design and characteristics. An interesting microphone of this design is the Schoeps Model CMT26U manufactured in West Germany by Schall-Technik, and named after Dr. Carl Schoeps, the designer.

The basic circuitry is shown in Fig. 16-48. By means of a single transistor, two oscillatory circuits are excited and tuned to the exact same frequency of 3.7 MHz. The output voltage from the circuits is rectified by a phase-bridge detector circuit, which operates over a large linear modulation range with very small RF voltages from the oscillator. The amplitude and polarity of the output voltage from the bridge depend on the phase angle between the two high-frequency voltages. The microphone capsule (head) acts as a variable capacitance in one of the oscillator circuits. When a sound wave impinges on the surface of the diaphragm of the microphone head, the vibrations of the diaphragm are detected by the phase curve of the oscillator circuit, and an audio frequency voltage is developed at the output of the bridge circuit. The microphone-head diaphragm is metal to guarantee a large constant capacitance. An automatic frequency control (afc) with a large range of operation is provided by means of capacitance diodes to preclude any influence caused by aging or temperature changes on the frequency-determining elements, that might throw the circuitry out of balance.

Internal output resistance is about 200 Ω. The signal, fed directly from the bridge circuit through two capacitors, delivers an output level of −51 dB to −49 dB (depending on the polar pattern used) into a 200 Ω load for a sound pressure level of 10 dynes/cm^2. The SNR and the distortion are independent of the load because of the bridge circuit; therefore, the microphone may be operated into load impedances ranging from 30 to 200 Ω.

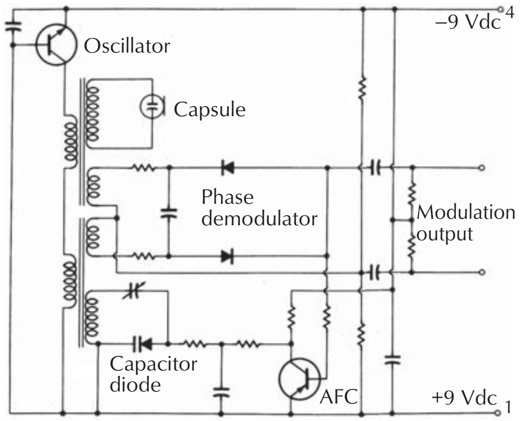


Figure 16-48. Basic circuit for the Schoeps radio-frequency capacitor microphone, series CMT.

16.3.4.3.1 Capacitor Radio-Frequency Microphones

A capacitor microphone of somewhat different design, manufactured by Sennheiser and also employing a crystal-controlled oscillator, is shown in Fig. 16-49. In the conventional capacitor microphone (without an oscillator) the input impedance of the preamplifier is in the order of 100 MΩ so it is necessary to place the capacitor head and preamplifier in close proximity. In the Sennheiser microphone, the capacitive element (head) used with the RF circuitry is a much lower impedance since the effect of a small change in capacitance at radio frequencies is considerably greater than at audio frequencies. Instead of the capacitor head being subjected to a high dc polarizing potential, the head is subjected to an RF voltage of only a few volts. An external power supply of 12 Vdc is required.

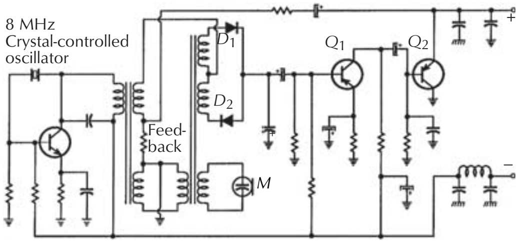


Figure 16-49. Basic circuit for the Sennheiser model 105, 405, and 805 capacitor microphones. Courtesy Sennheiser Electronic Corporation.

Referring to Fig. 16-49, the output voltage of the 8 MHz oscillator is periodically switched by diodes D_1 and D_2 to capacitor C. The switching phase is shifted 90° from that of the oscillator by means of loose coupling and individually aligning the resonance of the microphone circuit M under a no-sound condition. As a result, the voltage across capacitor C is zero. When a sound impinges on the diaphragm, the switching phase changes proportionally to the sound pressure, and a corresponding audio voltage appears across capacitor C. The output of the switching diodes is directly connected to the transistor amplifier stage, whose gain is limited to 12 dB by the use of negative feedback.

A high Q oscillator circuit is used to eliminate the effects of RF oscillator noise as noise in an oscillatory circuit is inversely proportional to the Q of the circuit. Because of the high Q of the crystal and its stability, compensating circuits are not required, resulting in low internal noise.

The output stage is actually an impedance-matching transformer adjusted for 100 Ω, for a load impedance of 2000 Ω or greater. RF chokes are connected in the output circuit to prevent RF interference and also to prevent external RF fields from being induced into the microphone circuitry.

16.3.4.3.2 Symmetrical Push-Pull Transducer Microphone

Investigations on the linearity of condenser microphones customarily used in the recording studios was carried out by Sennheiser using the *difference frequency method* incorporating a twin tone signal, Fig. 16-50. This is a very reliable test method as the harmonic distortions of both loudspeakers that generate the test sounds separately do not disturb the test result. Thus, difference frequency signals arising at the microphone output are arising from nonlinearities of the microphone itself.

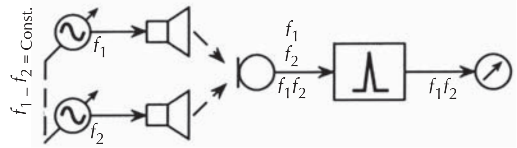


Figure 16-50. Difference frequency test. Courtesy Sennheiser Electronic Corporation.

Fig. 16-51 shows the distortion characteristics of eight unidirectional studio condenser microphones which were stimulated by two sounds of 104 dB SPL (3 Pa). The frequency difference was fixed to 70 Hz while the twin tone signal was swept through the upper audio range. The curves show that unwanted difference frequency signals of considerable levels were generated by all examined microphones. Although the curves are shaped rather individually, there is a general tendency for increased distortion levels (up to 1% and more) at high frequencies.

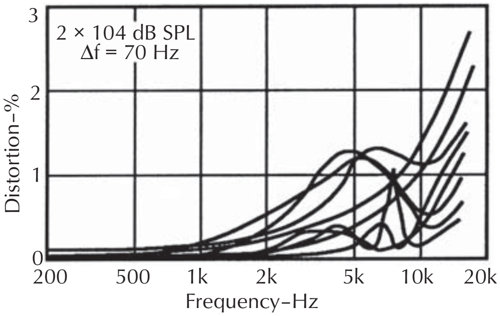


Figure 16-51. Frequency distortion of eight unidirectional microphones. Courtesy Sennheiser Electronic Corporation.

The measurement results can be extended to higher signal levels simply by linear extrapolation. This means, for instance, that 10 times higher sound pressures will yield 10 times higher distortions, as long as clipping of the microphone circuit is prevented. Thus, two sounds of 124 dB SPL will cause more than 10% distortion in the microphones. Sound pressure levels of this order are beyond the threshold of pain of human hearing but may arise at close-up micing. Despite the fact that the audibility of distortions depends significantly on the tonal structure of the sound signals, distortion figures of this order will considerably affect the fidelity of the sound pick-up.

The Cause of Nonlinearity. Fig. 16-52 shows a simplified sketch of a capacitive transducer. The diaphragm and backplate form a capacitor, the capacity of which depends on the width of the air gap. From the acoustical point of view the air gap acts as a complex impedance. This impedance is not constant but depends on the actual position of the diaphragm. Its value is increased if the diaphragm is moved toward the backplate and it is decreased at the opposite movement, so the air gap impedance is varied by the motion of the diaphragm. This implies a parasitic rectifying effect superimposed to the flow of volume velocity through the transducer, resulting in nonlinearity-created distortion.

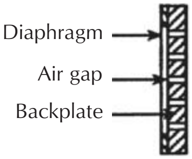


Figure 16-52. Conventional capacitor microphone transducer.

Solving the Linearity Problem. A push-pull design of the transducer helps to improve the linearity of condenser microphones, Fig. 16-53. An additional plate equal to the backplate is positioned symmetrically in front of the diaphragm, so two air gaps are formed with equal acoustical impedances as long as the diaphragm is in its rest position. If the diaphragm is deflected by the sound signal, then both air gap impedances are deviated opposite to each other. The impedance of one side increases while the other impedance decreases. The variation effects compensate each other regardless of the direction of the diaphragm motion, and the total air gap impedance is kept constant, reducing the distortion of a capacitive transducer.

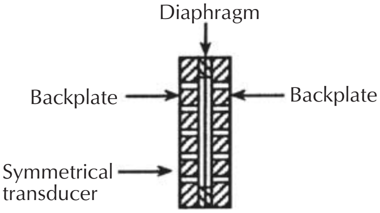


Figure 16-53. Symmetrical capacitor microphone transducer.

Fig. 16-54 shows the distortion characteristics of the Sennheiser MKH series push-pull element transformerless RF condenser microphones. The improvement on linearity due to the push-pull design can be seen by comparing Fig. 16-51 to Fig. 16-54.

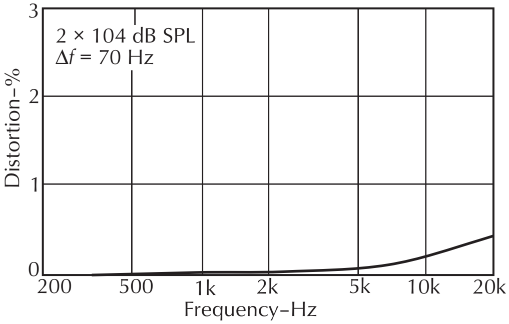


Figure 16-54. Distortion characteristics of the symmetrical capacitor microphone transducer.

16.3.4.3.3 Noise Sources

The inherent noise of condenser microphones is caused partly by the random incidence of the air particles at the diaphragm due to their thermal movement. The laws of statistics imply that sound pressure signals at the diaphragm can be evaluated by a precision that improves linearly with the diameter of the diaphragm. Thus, larger diaphragms yield better noise performance than smaller ones.

Another contribution of noise is the frictional effects in the resistive damping elements of the transducer. The noise generation from acoustical resistors is based on the same principles as the noise caused by electrical resistors so high acoustical damping implies more noise than low damping.

Noise is also added by the electrical circuit of the microphone. This noise contribution depends on the sensitivity of the transducer. High transducer sensitivity reduces the influence of the circuit noise. The inherent noise of the circuit itself depends on the operation principle and on the technical quality of the electrical devices.

Noise Reduction. Large-diameter diaphragms improve noise performance. Unfortunately, a large diameter increases the directivity at high frequencies. A 1 inch (25 mm) transducer diameter is usually a good choice.

A further method to improve the noise characteristics is the reduction of the resistive damping of the transducer. In most directional condenser microphones, a high amount of resistive damping is used in order to realize a flat frequency response of the transducer itself. With this design the electrical circuit of the microphone is rather simple. However, it creates reduced sensitivity and increased noise.

Keeping the resistive damping of the transducer moderate will be a more appropriate method to improve noise performance, however it leads to the transducer frequency response that is not flat so equalization has to be applied by electrical means to produce a flat frequency response of the complete microphone. This design technique requires a more sophisticated electrical circuit but produces good noise performance.

The electrical output of a transducer acts as a pure capacitance. Its impedance decreases as the frequency increases so the transducer impedance is low in an RF circuit but high in an AF circuit. Moreover, in an RF circuit the electrical impedance of the transducer does not depend on the actual audio frequency but is rather constant due to the fixed frequency of the RF oscillator. Contrary to this, in an AF design, the transducer impedance depends on the actual audio frequency, yielding very high values especially at low frequencies. Resistors of extremely high values are needed at the circuit input to prevent loading of the transducer output. These resistors are responsible for additional noise contribution.

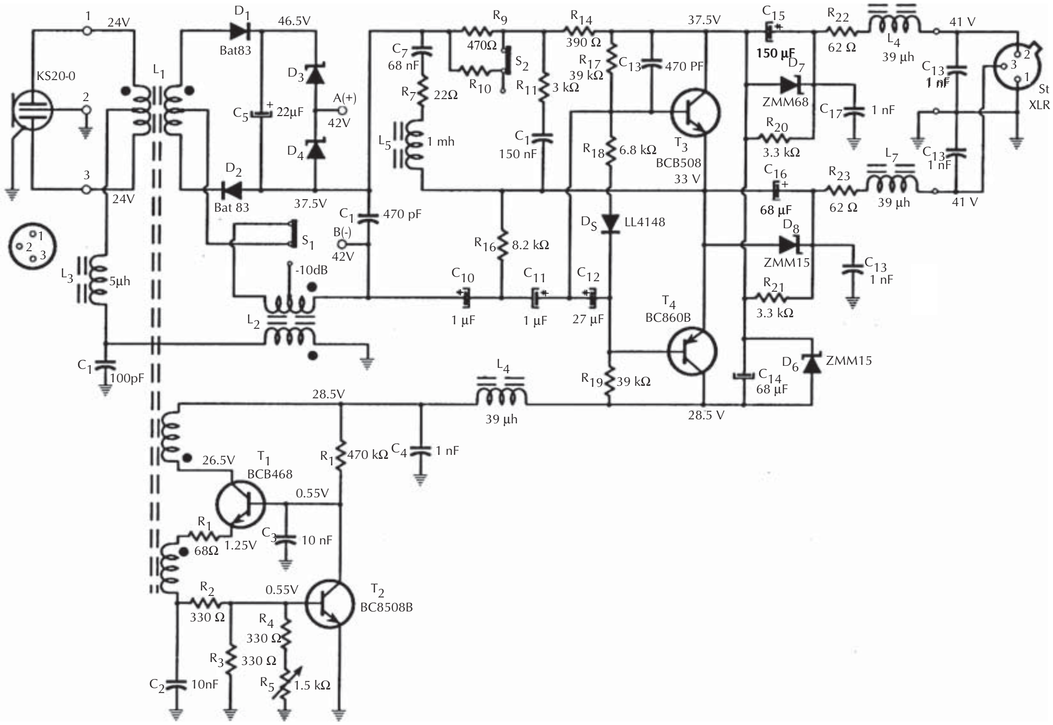
The RF circuit features a very low output impedance which is comparable to that of dynamic-type microphones. The output signal can be applied directly to bipolar transistors, yielding low noise performance by impedance matching.

The Sennheiser MKH 20, Fig. 16-55, is a pressure microphone with omnidirectional characteristics. The MKH 30 is a pure pressure-gradient microphone with a highly symmetrical bidirectional pattern due to the symmetry of the push-pull transducer. The MKH 40, Fig. 16-56, operates as a combined pressure and pressure-gradient microphone yielding a unidirectional cardioid pattern.

Figure 16-55. Omnidirectional pressure capacitor microphone. Note the lack of rear entry holes in the case. Courtesy Sennheiser Electronic Corporation.

Figure 16-56. Unidirectional pressure/pressure-gradient capacitor microphone. Courtesy Sennheiser Electronic Corporation.

- The microphones are phantom powered by 48 Vdc and 2 mA. The outputs are transformerless floated, Fig. 16-57.
- The SPL_{max} is 134 dB at nominal sensitivity and 142 dB at reduced sensitivity.
- The equivalent SPL of the microphones range from 10–12 dBA corresponding to CCIR-weighted figures of 20–22 dB.
- The directional microphones incorporate a switchable bass roll-off to cancel the proximity effect at close-up micing. The compensation is adjusted to about 5 cm (2 in) distance.



KS20-0
24V
Bat83
46.5V
Bat 83
37.5V
470 pF
470Ω
22Ω
68 nF
1 mh
150 nF
390 Ω
39 kΩ
3 kΩ
470 PF
6.8 kΩ
BCB508
33 V
LL4148
8.2 kΩ
1 µF
27 µF
BC860B
39 kΩ
28.5 V
150 µF
ZMM68
3.3 kΩ
68 µF
ZMM15
62 Ω
39 µh
1 nF
41 V
XLR 3
ZMM15
68 µF
-10dB
42V
5µh
100pF
28.5V
39 µh
26.5V
BCB468
470 kΩ
0.55V
68Ω
1.25V
10 nF
330 Ω
BC8508B
1.5 kΩ
10nF

Figure 16-57. Schematic of a Sennheiser MKH 20 P 48 U 3 capacitor microphone. Courtesy Sennheiser Electronic Corporation.

A special feature of the omnidirectional microphone is a switchable diffuse field correction that corrects for both direct and diffuse sound field conditions. The normal switch position is recommended for a neutral pickup when closeup micing and the diffuse field position is used if larger recording distances are used where reverberations become significant.

The distinction between both recording situations arises because omnidirectional microphones tend to attenuate lateral and reverse impinging sound signals at high frequencies. Diffuse sound signals with random incidence cause a lack of treble response, which can be compensated by treble emphasis at the microphone. Unfortunately, frontally impinging sounds are emphasized also, but this effect is negligible if the reverberant sound is dominant.

16.3.5 Electret Microphones

An *electret microphone* is a capacitor microphone in which the head capacitor is permanently charged, eliminating the need for a high-voltage bias supply.

From a design viewpoint a microphone intended to be used for critical recording, broadcast, or sound reinforcement represents a challenge involving minimal performance compromise. Early electrets offered the microphone designer a means of reducing the complexity of a condenser microphone by eliminating the high-voltage bias supply, but serious environmental stability problems negated this advantage. Well-designed electret microphones can be stored at 50°C (122°F) and 95% relative humidity for years with a sensitivity loss of only 1 dB. Under normal conditions of temperature and humidity, electret transducers will demonstrate a much lower charge reduction versus time than under the severe conditions indicated. Even if a proper electret material is used, there are many steps in the fabricating, cleaning, and charging processes that greatly influence charge stability.

The Shure SM81 cardioid condenser microphone[3] Fig. 16-58, uses an electret material as a means of establishing a bias voltage on the transducer. The backplate carries the electret material based upon the physical properties of halocarbon materials such as Teflon™ and Aclar, which are excellent electrets, and materials such as polypropylene and polyester terephthalate (Mylar™), which are more suitable for diaphragms.

Figure 16-58. Shure SM81 electret capacitor microphone. Courtesy Shure Incorporated.

The operation of the Shure SM-81 microphone is explained in Section 16.2.3.4.

16.4 Microphone Sensitivity[7]

Microphone sensitivity is the measure of the electrical output of a microphone with respect to the acoustic sound pressure level input.

Sensitivity is measured in one of three methods:

Open-circuit voltage	0 dB = 1 V/µbar
Maximum power output	0 dB = 1 mW/10 µbar = 1 mW/Pa
Electronic Industries Association (EIA) sensitivity	0 dB = EIA standard SE-105

The common sound pressure levels used for measuring microphone sensitivity are:

94 dB SPL	10 dyn/cm² SPL	10 µbar or 1 Pa
74 dB SPL	1 dyn/cm² SPL	1 µbar or 0.1 Pa
0 dB SPL	0.0002 dyn/cm² SPL	0.0002 Pa or 20 µPa—threshold of hearing

94 dB SPL is recommended since 74 dB SPL is too close to typical noise levels.

16.4.1 Open-Circuit Voltage Sensitivity

There are several good reasons for measuring the open-circuit voltage:

- If the open-circuit voltage and the microphone impedance are known, the microphone performance can be calculated for any condition of loading.
- It corresponds to an effective condition of use. A microphone should be connected to a high impedance to yield maximum SNR. A 150–250 Ω microphone should be connected to 2 kΩ or greater.
- When the microphone is connected to a high impedance compared to its own, variations in microphone impedance do not cause variations in response.

The open-circuit voltage sensitivity (S_v) can be calculated by exposing the microphone to a known SPL, measuring the voltage output, and using the following equation:

$$S_v = 20\log E_o - dB_{SPL} + 94 \qquad (16\text{-}4)$$

where,

S_v is the open-circuit voltage sensitivity in decibels re 1 V for a 10 dyn/cm² SPL (94 dB SPL) acoustic input to the microphone,

E_o is the output of the microphone in volts,

dB_{SPL} is the level of the actual acoustic input.

The microphone measurement system can be setup as shown in Fig. 16-59. The setup requires a random-noise generator, a microvoltmeter, a high-pass and a low-pass filter set, a power amplifier, a test-loudspeaker, and a sound level meter (SLM). The SLM is placed a specific measuring distance (about 5–6 ft or 1.5–2 m) in front of the loudspeaker. The system is adjusted until the SLM reads 94 dB SPL (a band of pink noise from 250 to 5000 Hz is excellent for this purpose). The microphone to be tested is now substituted for the SLM.

It is often necessary to know the voltage output of the microphone for various SPLs to determine whether the microphone will overload the preamplifier circuit or the SNR will be inadequate. To determine this, use

$$E_o = 10^{\left(\frac{S_v + dB_{SPL} - 94}{20}\right)} \qquad (16\text{-}5)$$

where,

E_o is the voltage output of microphone,

S_v is the open-circuit voltage sensitivity,

dB_{SPL} is the sound pressure level at the microphone.

16.4.2 Maximum Power Output Sensitivity[7]

The *maximum power output sensitivity* form of specification gives the maximum power output in decibels available from the microphone for a given sound pressure and power reference. Such a specification can be

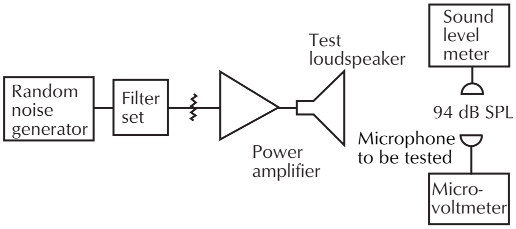


Figure 16-59. Method of determining open-circuit voltage sensitivity of a microphone. (From Reference 7.)

calculated from the internal impedance and the open-circuit voltage of the microphone. This specification also indicates the ability of a microphone to convert sound energy into electrical power. The equation is

$$S_p = 10\log\frac{V_o^2}{R_o} + 44\ \text{dB} \qquad (16\text{-}6)$$

where,

S_p is the power level microphone sensitivity in decibels,

V_o is the open-circuit voltage produced by a 1 μbar (0.1 Pa) sound pressure,

R_o is the internal impedance of the microphone.

The form of this specification is similar to the voltage specification except that a power as opposed to a voltage reference is given with the sound pressure reference. A 1 mW power reference and a 10 μbar (1 Pa) pressure reference are commonly used (as for the previous case). This form of microphone specification is quite meaningful because it takes into account both the voltage output and the internal impedance of the microphone.

S_p can also be calculated easily from the open-circuit voltage sensitivity

$$S_p = S_v - 10\log Z + 44\ \text{dB} \qquad (16\text{-}7)$$

where,

S_p is the decibel rating for an acoustical input of 94 dB_{SPL} (10 dyn/cm^2) or 1 Pa,

Z is the measured impedance of the microphone (the specifications of most manufacturers use the rated impedance).

The output level can also be determined directly from the open-circuit voltage

$$S_p = 10\log\frac{E_o^2}{0.001Z} - 6\ \text{dB} \qquad (16\text{-}8)$$

where,

E_o is the open-circuit voltage,

Z is the microphone impedance.

Because the quantity $10\log(E^2/0.001Z)$ treats the open-circuit voltage as if it appears across a load, it is necessary to subtract 6 dB. (The reading is 6 dB higher than it would have been had a load been present.)

16.4.3 Electronic Industries Association (EIA) Output Sensitivity

The Electronic Industries Association (EIA) Standard SE-105, August 1949, defines the system rating (G_M) as the ratio of the maximum electrical output from the microphone to the square of the undisturbed sound field pressure in a plane progressive wave at the microphone in decibels relative to 1 mW/0.0002 dyn/cm^2. Expressed mathematically,

$$G_M = 20\log\frac{E_o}{P} - 10\log Z_o - 50\ \text{dB} \qquad (16\text{-}9)$$

where,

E_o is the open-circuit voltage of the microphone,

P is the undisturbed sound field pressure in dyn/cm^2,

Z_o is the microphone-rated output impedance in ohms.

For all practical purposes, the output level of the microphone can be obtained by adding the sound pressure level relative to 0.0002 dyn/cm^2 to G_M.

Because G_M, S_V, and S_P are compatible, G_M can also be calculated

$$G_M = S_v - 10\log R_{MR} - 50\ \text{dB} \qquad (16\text{-}10)$$

where,

G_M is the EIA rating,

R_{MR} is the EIA center value of the nominal impedance range shown below.

Ranges (ohms)		Values Used (ohms)
20–80	=	38
80–300	=	150
300–1250	=	600
1250–4500	=	2400
4500–20,000	=	9600
20,000–70,000	=	40,000

The EIA rating can also be determined from the chart in Fig. 16-60.

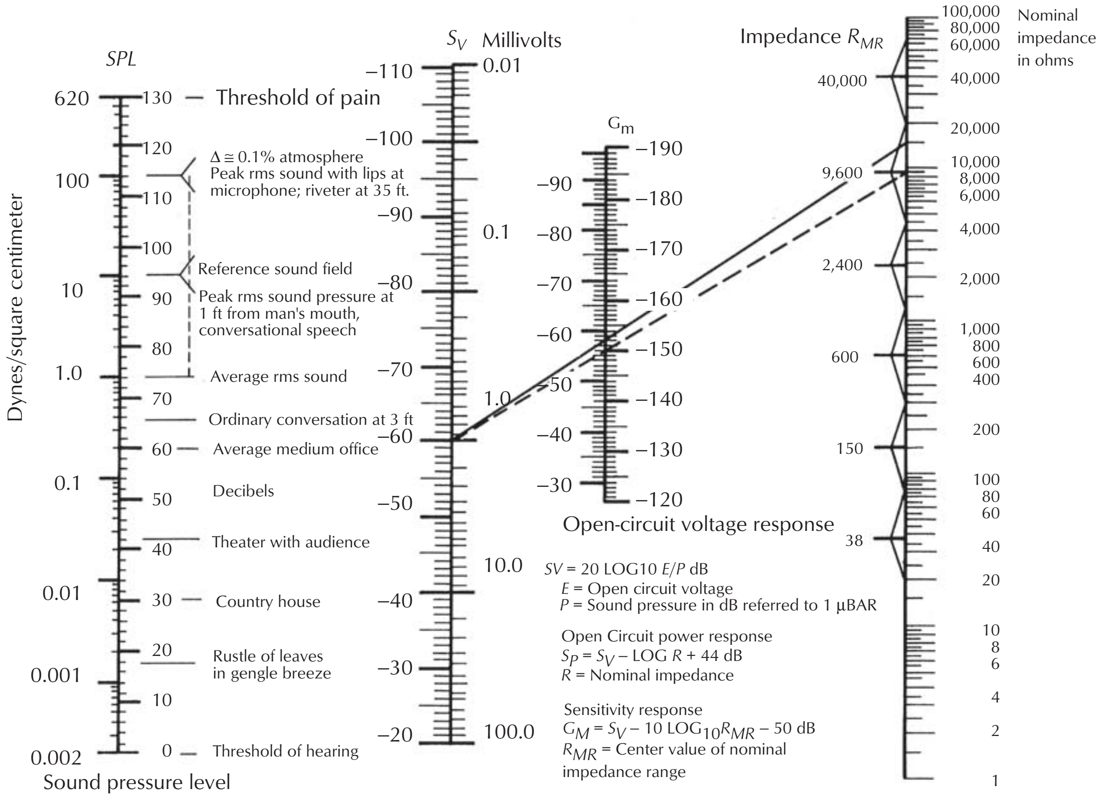


Figure 16-60. Microphone sensitivity conversion chart.

16.4.4 Various Microphone Sensitivities

Microphones are subjected to sound pressure levels anywhere from 40 dB SPL when distant micing to 150 dB SPL when extremely close micing (i.e., ¼ inch from the rock singer's mouth or inside a drum or horn).

Various types of microphones have different sensitivities, which is important to know if different types of microphones are intermixed since gain settings, SNR, and preamplifier overload will vary. Table 16-1 gives the sensitivities of a variety of different types of microphones.

16.4.5 Microphone Thermal Noise

Since a microphone has an impedance, it generates thermal noise. Even without an acoustic signal, the microphone will still produce a minute output voltage. The thermal noise voltage, E_n, produced by the electrical resistance of a sound source is dependent on the frequency bandwidth under consideration, the magnitude of the resistance, and the temperature existing at the time of the measurement. This voltage is

$$E_n = 4ktR(bw) \qquad (16\text{-}11)$$

where,
k is the Boltzmann's constant, 1.38×10^{-23} J/K,
t is the absolute temperature, 273° + room temperature, both in °C,
R is the resistance in ohms,
bw is the bandwidth in hertz.

To change this to *dBv* use

$$EIN_{dBv} = 20\log\frac{E_n}{0.775} \qquad (16\text{-}12)$$

The thermal noise relative to 1 V is −198 dB for a 1 Hz bandwidth and 1 Ω impedance. Therefore,

$$\frac{TN}{1\text{ V}} = -198\text{ dB} + 10\log(bw) + 10\log Z \qquad (16\text{-}13)$$

where,

TN is the thermal noise relative to 1 V,
bw is the bandwidth in hertz,
Z is the microphone impedance in ohms.

Thermal noise relative to 1 V can be converted to equivalent input noise (EIN) by

$$EIN_{dBm} = -198\text{ dB} + 10\log(bw) + 10\log Z - 6 - 20\log 0.775\text{ V}. \tag{16-14}$$

Since the *EIN* is in dBm and dBm is referenced to 600 Ω, the impedance Z is 600 Ω.

Table 16-1. Sensitivities of Various Types of Microphones

Type of Microphone	Sp	Sv
Carbon-button	–60 to –50 dB	
Crystal		–50 to –40 dB
Ceramic		–50 to –40 dB
Dynamic (moving coil)	–60 to –52 dB	–85 to –70 dB
Capacitor	–60 to –37 dB	–85 to –45 dB
Ribbon-velocity	–60 to –50 dB	–85 to –70 dB
Transistor	–60 to –40 dB	
Sound power	–32 to –20 dB	
Line level	–40 to 0 dB	–20 to 0 dB
Wireless	–60 to 0 dB	–85 to 0 dB

16.5 Microphone Practices

16.5.1 Placement

Microphones are placed in various relationships to the sound source to obtain various sounds. Whatever position gives the desired effect that is wanted is the correct position. There are no exact rules that must be followed, however, certain recommendations should be followed to assure a good sound.

16.5.1.1 Microphone-to-Source Distance

Microphones are normally used in the direct field. Under this condition, inverse square law attenuation prevails, meaning that each time the distance is doubled, the microphone output is reduced 6 dB. For instance, moving from a microphone-to-source distance of 2.5 to 5 cm (1 to 2 in) has the same effect as moving from 15 to 30 cm (6 to 12 in), 1 to 2 ft (30 to 60 cm), or 5 to 10 ft (1.5 to 3 m).

Distance has many effects on the system. In a reinforcement system, doubling the distance reduces gain before feedback 6 dB; in all systems, it reduces the effect of microphone-to-source variations.

Using the inverse-square-law equation for attenuation,

$$attenuation_{dB} = 20\log\frac{D_1}{D_2} \tag{47-15}$$

it can be seen, at a microphone-to-source distance of 2.5 cm (1 in), moving the microphone only 1.25 cm (½ in) closer will increase the signal 6 dB and 1.25 cm (½ in) farther away will decrease the signal 3.5 dB for a total signal variation of 9.5 dB for only 2.5 cm (1 in) of total movement! At a source-to-microphone distance of 30 cm (12 in), a movement of 2.5 cm (1 in) will cause a signal variation of only 0.72 dB. Both conditions can be used advantageously; for instance, close micing is useful in feedback-prone areas, high noise level areas (rock groups), or where the talent wants to use the source to microphone variations to create an effect.

The farther distances are most useful where lecterns and table microphones are used or where the talker wants movement without level change.

The microphone-to-source distance also has an effect on the sound of a microphone, particularly one with a cardioid pattern. As the distance decreases, the proximity effect increases creating a bassy sound (see Section 16.2.3.1). Closing in on the microphone also increases breath noise and pop noise.

16.5.1.2 Distance from Large Surfaces

When a microphone is placed next to a large surface such as the floor, 6 dB of gain can be realized, which can be a help when far micing.

As the microphone is moved away from the large surface but still in proximity of it, cancellation of some specific frequencies will occur, creating a notch of up to 30 dB, Fig. 16-65. The notch is created by the cancellation of a frequency that, after reflecting off the surface, reaches the microphone diaphragm 180° out of polarity with the direct sound.

The frequency of cancellation, f_c, can be calculated from the equation

$$f_c = \frac{0.5c}{D_{r1} + D_{r2} - D_d} \tag{16-16}$$

where,
c is the speed of sound, 1130 feet per second or 344 meters per second,
0.5 is the out-of-polarity frequency ratio,
D_{r1} is the reflected path from the source to the surface in feet or meters,

D_{r2} is the reflected path from the surface to the microphone in feet or meters,

D_d is the direct path from the source to the microphone in feet or meters.

If the microphone is 10 ft from the source and both are 5 ft above the floor, the canceled frequency is

$$
\begin{aligned}
f_c &= \frac{1130 \times 0.5}{7.07 + 7.07 - 10} \qquad (16\text{-}17)\\
&= 136.47\ \text{Hz}
\end{aligned}
$$

If the microphone is moved to 2 ft above the floor, the canceled frequency is 319.20 Hz. If the microphone is 6 inches from the floor, the canceled frequency is 1266.6 Hz. If the microphone is 1 inch from the floor, the canceled frequency is 7239.7 Hz.

16.5.1.3 Behind Objects

Sound, like light, does not go through solid or acoustically opaque objects. It does, however, go through objects of various density. The transmission loss or ability of sound to go through this type of material is frequency dependent; therefore, if an object of this type is placed between the sound source and the microphone, the pickup will be attenuated according to the transmission characteristics of the object.

Low-frequency sound bends around objects smaller than their wavelength, which affects the frequency response of the signal. The normal effect of placing the microphone behind an object is an overall reduction of level, a low-frequency boost, and a high-frequency roll-off.

16.5.1.4 Above the Source

When the microphone is placed above or to the side of a directional sound source (i.e., horn or trumpet), the high-end frequency response will roll off because high frequencies are more directional than low frequencies, so less high-frequency SPL will reach the microphone than low-frequency SPL.

16.5.1.5 Direct versus Reverberant Field

Micing in the reverberant field picks up the characteristic of the room because the microphone is picking up as much or more of the room, as it is the direct sound from the source. When micing in the reverberant field, only two microphones are required for stereo since isolation of the individual sound sources is impossible. When in the reverberant field, a directional microphone will lose much of its directivity. Therefore, it is often advantageous to use an omnidirectional microphone that has smoother frequency response. To mic sources individually, you must be in the direct field and usually very close to the source to eliminate cross-feed.

16.5.2 Grounding

The grounding of microphones and their interconnecting cables is of extreme importance since any hum or noise picked up by the cables will be amplified along with the audio signal. Professional systems generally use the method shown in Fig. 16-61. Here the signal is passed through a two-conductor shielded cable to the balanced input of a preamplifier. The cable shield is connected to pin number 1, and the audio signal is carried by the two conductors and pins 2 and 3 of the XLR-type connector. The actual physical ground is connected at the preamplifier chassis only and carried to the microphone case. In no instance is a second ground ever connected to the far end of the cable, because this will cause the flow of ground currents between two points of grounding.

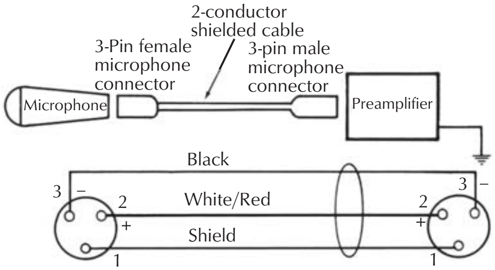

2-conductor shielded cable
3-Pin female microphone connector
3-pin male microphone connector
Microphone
Preamplifier
Black
3
–
2
+
White/Red
Shield
1
3
–
2
+
1

Figure 16-61. Typical low-impedance microphone to preamplifier wiring.

In systems designed for semiprofessional and home use, the method in Fig. 16-62 is often used. Note that one side of the audio signal is carried over the cable shield to a pin-type connector. The bodies of both the male and female connector are grounded: the female to the amplifier case and the male to the cable shield. The microphone end is connected in a similar manner; here again the physical ground is connected only at the preamplifier chassis. Hum picked up on the shield and not on the center conductor is added to the signal and amplified through the system.

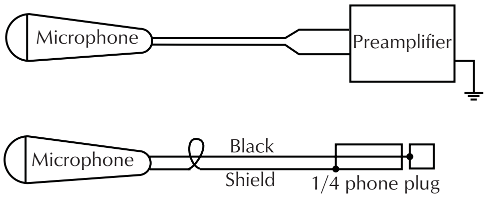

Microphone
Preamplifier
Microphone
Black
Shield
1/4 phone plug

Figure 16-62. Typical semiprofessional, hi-fi microphone to preamplifier wiring.

16.5.3 Polarity

Microphone polarity, or *phase* as it is often called, is important especially when multiple microphones are used. When they are in polarity they add to each other rather than have canceling effects. If multiple microphones are used and one is out of polarity, it will cause comb filters, reducing quality and stereo enhancement. The EIA standard RS-221.A, October 1979, states "Polarity of a microphone or a microphone transducer element refers to in-phase or out-of-phase condition of voltage developed at its terminals with respect to the sound pressure of a sound wave causing the voltage."

Note: *Exact in-phase relationship* can be taken to mean that the voltage is coincident with the phase of the sound pressure wave causing the voltage. In practical microphones, this perfect relationship may not always be obtainable.

The positive or in-phase terminal is that terminal that has a positive potential and a phase angle less than 90° with respect to a positive sound pressure at the front of the diaphragm.

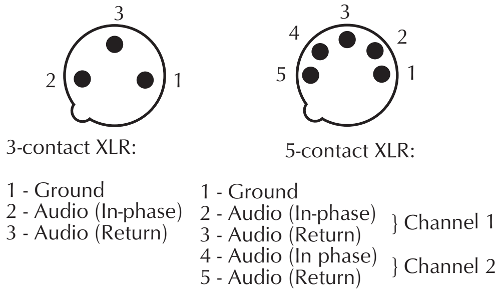
When connected to a *three-pin XLR connector* as per EIA standard RS-297, the polarity shall be as follows:

- Out-of-phase—terminal 3 (black).
- In-phase—terminal 2 (red or any color other than black).
- Ground—terminal 1 (shield).

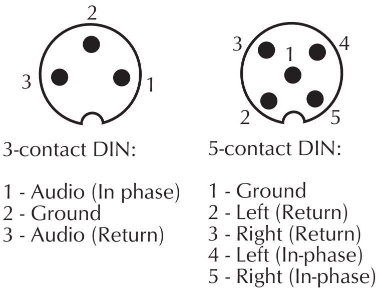
Fig. 16-63 shows the proper polarity for three-pin and five-pin XLR connectors and for three-pin and five-pin DIN connectors.

A simple method of determining microphone polarity is as follows:

If two microphones have the same frequency response and sensitivity and are placed next to each other and connected to the same mixer, the output will double if both are used. However, if they are out of polarity with each other, the total output will be down 40–50 dB from the output of only one microphone.


3
2
1
3-contact XLR:
1 - Ground
2 - Audio (In-phase)
3 - Audio (Return)
3
4
2
5
1
5-contact XLR:
1 - Ground
2 - Audio (In-phase)
3 - Audio (Return)
} Channel 1
4 - Audio (In phase)
5 - Audio (Return)
} Channel 2

A. XLR connectors.


2
3
1
3-contact DIN:
1 - Audio (In phase)
2 - Ground
3 - Audio (Return)
3
1
4
2
5
5-contact DIN:
1 - Ground
2 - Left (Return)
3 - Right (Return)
4 - Left (In-phase)
5 - Right (In-phase)

B. DIN connectors.

Figure 16-63. Microphone connector polarity.

The microphones to be tested for proper polarity are placed alongside each other and connected to their respective mixer inputs. With a single acoustic source into the microphones (pink noise is a good source), one mixer volume control is adjusted for a normal output level as indicated on a VU meter. Note the volume control setting and turn it off. Make the same adjustment for the second microphone, and note the setting of this volume control. Now open both controls to these settings. If the microphones are out of polarity, the quality of reproduction will be distorted, and there will be a distinct drop in level. Reversing the electrical connections to one microphone will bring them into polarity, making the quality about the same as one microphone operating and the output level higher.

If the microphones are of the bidirectional type, one may be turned 180° to bring it into polarity and later corrected electrically. If the microphones are of the directional type, only the output or cable connections can be reversed. After polarizing a bidirectional microphone, the rear should be marked with a white stripe for future reference.

16.5.4 Balanced or Unbalanced

Microphones can be connected either *balanced* or *unbalanced*. All professional installations use a balanced system for the following reasons:

- Reduced pickup of hum.

- Reduced pickup of electrical noise and transients.
- Reduced pickup of electrical signals from adjacent wires.

These reductions are realized because the two signal conductors shown in Fig. 16-63 pick up the same stray signal with equal intensity and polarity, so the noise is impressed evenly on each end of the transformer primary, eliminating a potential across the transformer and canceling any input noise. Because the balanced wires are in a shielded cable, the signal to each conductor is also greatly reduced.

When installing microphones into an unbalanced system, any noise that gets to the inner unbalanced conductor is not canceled by the noise in the shield, so the noise is transmitted into the preamplifier. In fact, noise impressed on the microphone end of the shield adds to the signal because of the resistance of the shield between the noise and the amplifier.

Balanced low-impedance microphone lines can be as long as 500 ft (150 m) but unbalanced microphone lines should never exceed 15 ft (4.5 m).

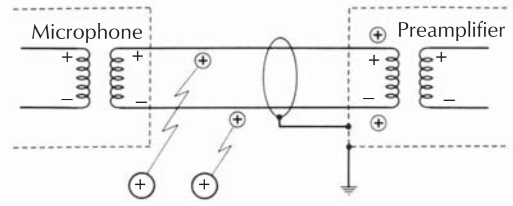


Figure 16-64. Noise cancellation on balanced, shielded microphone cables.

16.5.5 Impedance

Most professional microphones are low impedance, 200 Ω, and are designed to work into a load of 2000 Ω. High-impedance microphones are 50,000 Ω and are designed to work into an impedance of 1–10 MΩ. The low-impedance microphone has the following advantages:

- Less susceptible to noise. A noise source of relatively high impedance cannot “drive” into a source of relatively low impedance (i.e., the microphone cable).
- Capable of being connected to long microphone lines without noise pickup and high-frequency loss.

All microphone cable has inductance and capacitance. The capacitance is about 40 pF (40×10^{-12})/ft (30 cm). If a cable is 100 ft long (30 m), the capacitance would be (40×10^{-12}) × 100 ft or 4×10^{-9} F or 0.004 μF. This is equivalent to a 3978.9 Ω impedance at 10,000 Hz and is found with the equation

$$X_c = \frac{1}{2\pi fC} \tag{16-18}$$

This has little effect on a microphone with an impedance of 200 Ω as it does not reduce the impedance appreciably as determined by

$$Z_T = \frac{X_c Z_m}{X_c + Z_m} \tag{16-19}$$

For a microphone impedance of 200 Ω, the total impedance Z_T = 190 Ω or less than 0.5 dB.

If this same cable were used with a high-impedance microphone of 50,000 Ω, 10,000 Hz would be down more than 20 dB.

Making the load impedance equal to the microphone impedance will reduce the microphone sensitivity 6 dB, which reduces the overall SNR by 6 dB. For the best SNR, the input impedance of low-impedance microphone preamplifiers is always 2000 Ω or greater.

If the load impedance is reduced to less than the microphone impedance, or the load impedance is not resistive, the microphone frequency response and output voltage will be affected.

Changing the load of a high-impedance or ceramic microphone from 10 MΩ to 100 kΩ reduces the output at 100 Hz by 27 dB.

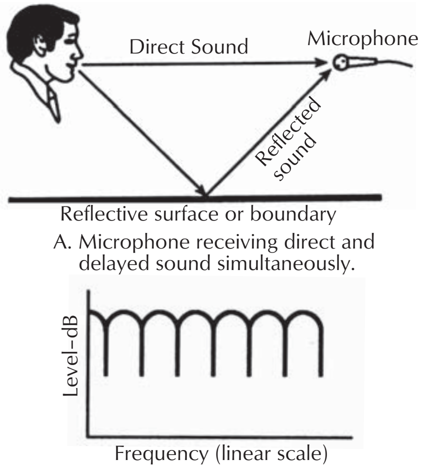
16.6 Miscellaneous Microphones

16.6.1 Pressure Zone Microphones (PZM)

The *pressure zone microphone*, referred to as a *PZMicrophone* or *PZM*, is a miniature condenser microphone mounted face-down next to a sound-reflecting plate or boundary. The microphone diaphragm is placed in the pressure zone just above the boundary where direct and reflected sounds combine effectively in-phase over the audible range.

In many recording and reinforcement applications, the sound engineer is forced to place microphones near hard reflective surfaces such as when recording an instrument surrounded by reflective baffles, reinforcing drama or opera with the microphones near the stage floor, or recording a piano with the microphone close to the open lid.

In these situations, sound travels from the source to the microphone via two paths: directly from the source to the microphone, and reflected off the surface to the microphone. The delayed sound reflections combine with the direct sound at the microphone, resulting in phase cancellations of various frequencies, Fig. 16-65. This creates a series of peaks and dips in the net frequency response called the *comb-filter effect*, affecting the recorded tone quality and giving an unnatural sound.



A. Microphone receiving direct and delayed sound simultaneously.

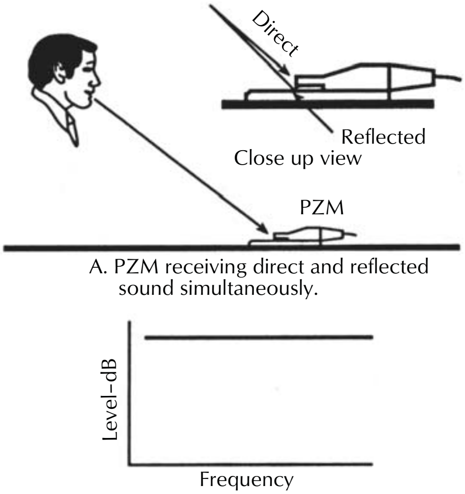
B. Resulting frequency response.

Figure 16-65. Effects of cancellation caused by near reflections (comb filters).

The PZM was developed to avoid the tonal coloration caused by microphone placement near a surface. The microphone diaphragm is arranged parallel with and very close to the reflecting surface and facing it, so that the direct and reflected waves combine at the diaphragm in-phase over the audible range, Fig. 16-66.

This arrangement can provide several benefits:

- Wide, smooth frequency response (natural reproduction) because of the lack of phase interference between direct and reflected sound.
- A 6 dB increase in sensitivity because of the coherent addition of direct and reflected sound.
- High SNR created by the PZM's high sensitivity and low internal noise.
- A 3 dB reduction in pickup of the reverberant field compared to a conventional omnidirectional microphone.
- Lack of off-axis coloration as a result of the sound entry's small size and radial symmetry.
- Good-sounding pickup of off-mic instruments due to the lack of off-axis coloration.
- Identical frequency response for random-incidence sound (ambience) and direct sound due to the lack of off-axis coloration.
- Consistent tone quality regardless of sound-source movement or microphone-to-source distance.
- Excellent reach (clear pickup of quiet distant sounds).
- Hemispherical polar pattern, equal sensitivity to sounds coming from any direction above the surface plane.
- Inconspicuous low-profile mounting.



A. PZM receiving direct and reflected sound simultaneously.

B. Resulting frequency response.

Figure 16-66. Effects of receiving direct and reflected sound simultaneously.

The Crown PZM-30 series microphones, Fig. 16-67, one of the original PZMs, are designed for professional use and built to take the normal abuse associated with professional applications. Miniaturized electronics built into the microphone cantilever allow the PZM-30 series to be powered directly by simplex phantom powering.

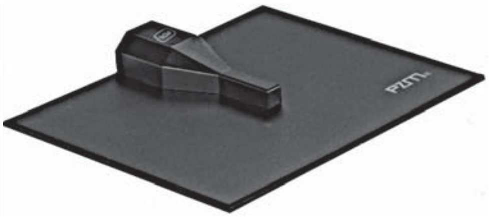


Figure 16-67. PZM 30D pressure zone microphone. Courtesy Crown International, Inc.

16.6.1.1 Phase Coherent Cardioid (PCC)

The *phase coherent cardioid microphone* (PCC) is a surface-mounted supercardioid microphone with many of the same benefits as the PZM. Unlike the PZM, however, the PCC uses a subminiature supercardioid microphone capsule.

Technically, the PCC is not a pressure zone microphone. The diaphragm of a PZM is parallel to the boundary; the diaphragm of the PCC is perpendicular to the boundary. Unlike a PZM, the PCC aims along the plane on which it is mounted. In other words, the main pickup axis is parallel with the plane.

The Crown PCC-160 microphone, a Phase Coherent Cardioid surface-mounted boundary microphone, Fig. 16-68, is intended for use on stage floors, lecterns, and conference tables wherever gain-before-feedback and articulation are important. Fig. 16-69 shows the horizontal polar response for this microphone.

Figure 16-68. Crown PCC®-200 Phase Coherent Cardioid microphone. Courtesy Crown International, Inc.

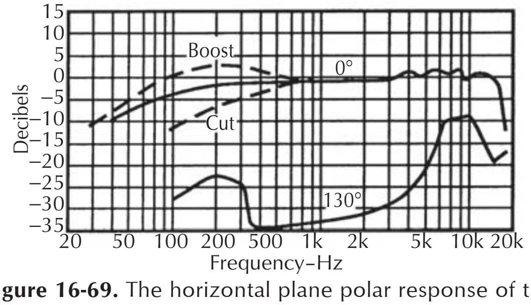


Figure 16-69. The horizontal plane polar response of the PCC-160 phase coherent cardioid microphone with the source 30° above the infinite boundary. Courtesy Crown International, Inc.

The PCC-160 can be directly phantom powered. A bass-tilt switch is provided for tailoring low end response.

16.6.1.2 Directivity

The PZM picks up sounds arriving from any direction above the surface it is mounted on (hemispherical). It is often necessary to discriminate against sounds arriving from certain directions. To make the microphone directional or hemicardioid (reject sounds from the rear) the capsule can be mounted with the cantilever in a corner boundary made of ¼ inch (6 mm) thick Plexiglas. The larger the boundary, the better it discriminates against low-frequency sounds from the rear.

For best results, a corner boundary 12 in × 24 in wide (0.3 m × 0.6 m) is recommended and is nearly invisible to the audience, Fig. 16-70.

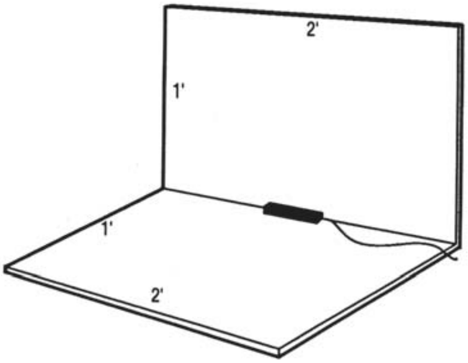


Figure 16-70. Corner boundary used to control directivity of pressure zone microphones.

A boom-mounted or suspended PZM can be taped to the center of a ¼ inch (6 mm) thick 2 ft × 2 ft (0.6 m × 0.6 m) or 4 ft × 4 ft (1.2 m × 1.2 m) panel. The microphone should be placed 4 inches (10 cm) off-center for a smoother frequency response. Using clear acrylic plastic (Plexiglas) makes the panel nearly invisible from a distance. If the edges of the Plexiglas pick up light, they can be taped or painted black.

16.6.1.2.1 Sensitivity Effects

If a PZM capsule is placed very near a single large boundary (within 0.020 inch or 0.50 mm), such as a large plate, floor, or wall, incoming sound reflects off the surface. The reflected sound wave adds to the incoming sound wave in the Pressure Zone next to the boundary. This coherent addition of sound waves doubles the sound pressure at the microphone, effectively increasing the microphone sensitivity or output by 6 dB over a standard microphone.

If the PZM capsule is placed at the junction of two boundaries at right angles to each other, such as the floor and a wall, the wall increases sensitivity 6 dB, and the floor increases sensitivity another 6 dB. Adding two boundaries at right angles increases sensitivity 12 dB.

With the PZM element at the junction of three boundaries at right angles, such as in the corner of the floor and two walls, microphone sensitivity will be 18 dB higher than what it was in open space.

Note that the acoustic sensitivity of the microphone rises as boundaries are added, but the electronic noise of the microphone stays constant, so the effective SNR of the microphone improves 6 dB every time a boundary is added at right angles to previous boundaries.

16.6.1.2.2 Direct-to-Reverberant Ratio Effects

Direct sound sensitivity increases 6 dB per boundary added, while reverberant or random-incidence sound increases only 3 dB per boundary added. Consequently, the direct-to-reverberant ratio increases 3 dB ($6\ \mathrm{dB}_{\mathrm{dir}} - 3\ \mathrm{dB}_{\mathrm{rev}}$) whenever a boundary is added at right angles to previous boundaries.

16.6.1.2.3 Frequency-Response Effects

The low-frequency response of the PZM or PCC depends on the size of the surface it is mounted on. The larger the surface, the more the low-frequency response is extended. The low-frequency response shelves down to a level 6 dB below the mid-frequency level at the frequency where the wavelength is about 6 times the boundary dimension. For example, the frequency response of a PZM on a 2 ft × 2 ft (0.6 m × 0.6 m) panel shelves down 6 dB below 94 Hz. On a 5 inch × 5 inch (12 cm × 12 cm) plate, the response shelves down 6 dB below 376 Hz.

For best bass and flattest frequency response, the PZM or PCC must be placed on a large hard boundary such as a floor, wall, table, or baffle at least 4 ft × 4 ft (1.2 m × 1.2 m).

To reduce bass response, the PZM or PCC can be mounted on a small plate well away from other reflecting surfaces. This plate can be made of thin plywood, Masonite, clear plastic, or any other hard, smooth material. When used on a carpeted floor the PZM or PCC should be placed on a hard-surfaced panel at least 1 ft × 1 ft (0.3 m × 0.3 m) for flattest high-frequency response.

To determine the frequency $f_{-6\ \mathrm{dB}}$ where the response shelves down 6 dB, use

$$f_{-6\ \mathrm{dB}} = \frac{188^*}{D} \tag{16-20}$$

*57.3 for SI units

where,

D is the boundary dimension in feet or meters.

For example, if the boundary is 2 ft (0.6 m) square, the 6 dB down point is

$$\begin{aligned} f_{-6dB} &= \frac{188}{D} \\ &= \frac{188}{2} \\ &= 94\ \mathrm{Hz} \end{aligned}$$

Below 94 Hz, the response is a constant 6 dB below the upper mid-frequency level. Note that there is a response shelf, not a continuous roll-off.

When the PZM is on a rectangular boundary, two shelves appear. The long side of the boundary is D_{max} and the short side D_{min}. The response is down 3 dB at

$$f_{-3dB} = \frac{188^*}{D_{max}} \tag{47-21}$$

*57.3 for SI units

and is down another 3 dB at

$$f_{-3dB} = \frac{188^*}{D_{min}} \tag{47-22}$$

*57.3 for SI units

The low-frequency shelf varies with the angle of the sound source around the boundary. At 90° incidence (sound wave motion parallel to the boundary), there is no low-frequency shelf.

The depth of the shelf also varies with the distance of the sound source to the panel. The shelf starts to disappear when the source is closer than a panel dimension away. If the source is very close to the PZM mounted on a panel, there is no low-frequency shelf; the frequency response is flat.

If the PZM is at the junction of two or more boundaries at right angles to each other, the response shelves down 6 dB per boundary at the above frequency. For example, a two-boundary unit made of 2 ft (0.6 m) square panels shelves down 12 dB below 94 Hz.

There are other frequency-response effects in addition to the low-frequency shelf. For sound sources on-axis to the boundary, the response rises about 10 dB above the shelf at the frequency where the wavelength equals the boundary dimension.

For a square panel,

$$F_{peak} = \frac{0.88c}{D} \tag{16-23}$$

where,

c is the speed of sound (1130 ft/s or 344 m/s)

D is the boundary dimension in feet or meters.

For a circular panel

$$F_{peak} = \frac{c}{D} \tag{16-24}$$

As an example, a 2 ft (0.6 m) square panel has a 10 dB rise above the shelf at

$$\begin{aligned} F_{peak} &= \frac{0.88c}{D} \\ &= \frac{0.88 \times 1130}{2} \\ &= 497\ \text{Hz} \end{aligned}$$

Note that this response peak is only for the direct sound of an on-axis source. The effect is much less if the sound field at the panel is partly reverberant, or if the sound waves strike the panel at an angle. The peak is also reduced if the microphone capsule is placed off-center on the boundary.

Fig. 16-71 shows the frequency response at various angles of sound incidence of a PZM mounted on a 2 ft square panel. Note the several phenomena shown in the figure:

- The low-frequency shelf (most visible at 30° and 60°).
- The lack of low-frequency shelving at 90° (grazing incidence).
- The 10 dB rise in response at 497 Hz.
- The edge-interference peaks and dips above 497 Hz (most visible at 0° or normal incidence).
- The lessening of interference at increasing angles.
- The greater rear rejection of high frequencies than low frequencies.

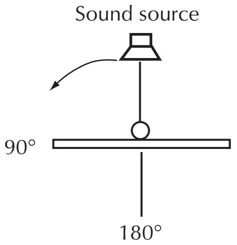
16.6.1.2.4 Frequency-Response Anomalies Caused by Boundaries

Frequency response is affected by:

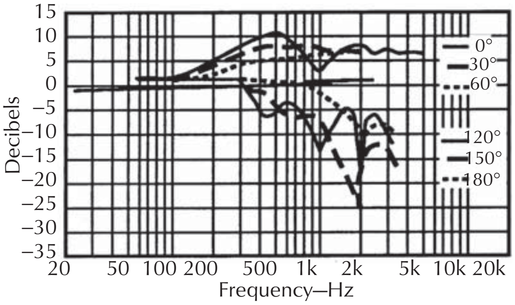
- When sound waves strike a boundary, pressure doubling occurs at the boundary surface, but does not occur outside the boundary, so there is a pressure difference at the edge of the boundary. This pressure difference creates sound waves.

  These sound waves generated at the edge of the boundary travel to the microphone in the center of the boundary. At low frequencies, these edge waves are opposite in polarity to the incoming sound waves. Consequently, the edge waves cancel the pressure-doubling effect.
- At low frequencies, pressure doubling does not occur, but at mid- to high-frequencies, pressure doubling does occur. The net effect is a mid- to high-frequency boost, which could be considered a low-frequency loss or shelf.
- Incoming waves having wavelengths about six times the boundary dimensions are canceled by edge effects while waves much smaller than the boundary dimensions are not canceled by edge effects.
- Waves having wavelengths on the order of the boundary dimensions are subject to varying interference versus frequency; i.e., peaks and dips in the frequency response.
- At the frequency where the wavelength equals the boundary dimension, the edge wave is in phase with the incoming wave. Consequently, there is a response rise (about 10 dB above the low frequency shelf) at that frequency. Above that frequency, there is a series

i



A. Layout.



B. Response.

Figure 16-71. Frequency response of a pressure zone microphone. Note the 6 dB shelf at 94 Hz.

of peaks and dips that decrease in amplitude with frequency.

- The edge-wave interference decreases if the incoming sound waves approach the boundary at an angle.
- Interference also is reduced by placing the microphone capsule off-center. This randomizes the distances from the edges to the microphone capsule, resulting in a smoother response.

16.6.2 Lavalier Microphones

Lavalier microphones are made either to wear on a lavalier around the neck or to clip onto a tie, shirt, or other piece of clothing. The older heavy style lavalier microphone, Fig. 16-72, which actually laid on the chest, had a frequency response that was shaped to reduce the boomingness of the chest cavity, and the loss of high-frequency response caused by being 90° off axis to the signal, Fig. 16-73. These microphones should not be used for anything except as a lavalier microphone.

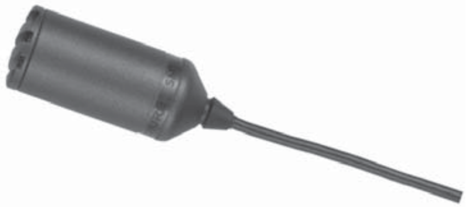


Figure 16-72. Shure SM11 dynamic omnidirectional lavalier microphone. Courtesy Shure Incorporated.

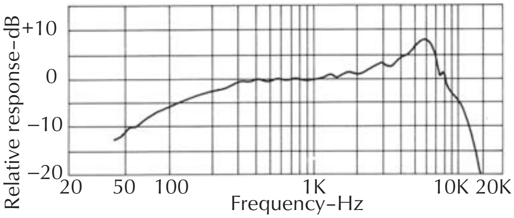


Figure 16-73. Typical frequency response of a heavy-style dynamic lavalier microphone.

Lavalier microphones may be dynamic, condenser (capacitor), pressure-zone, electret, or high-impedance ceramic.

The newer loss mass clip-on lavalier microphones, Figs. 16-74 and 16-75, do not require frequency response correction because there is no coupling to the chest cavity and the small diameter of the diaphragm does not create pressure build-up at high frequencies, creating directionality.

Most lavalier microphones are omnidirectional, however, more manufacturers are producing directional lavalier microphones. The Sennheiser MKE 104 clip-on lavalier microphone, Fig. 16-74, has a cardioid pickup pattern, Fig. 16-76. This reduces feedback, background noise, and comb filtering caused by the canceling between the direct sound waves and sound waves that hit the microphone on a reflective path from the floor, lectern, and so forth.

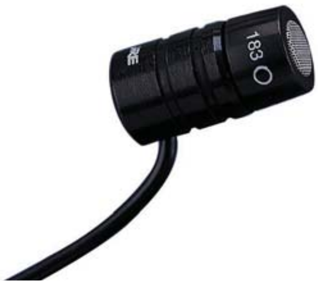


Figure 16-74. Shure SM183 omnidirectional condenser lavalier microphone. Courtesy Shure Incorporated.

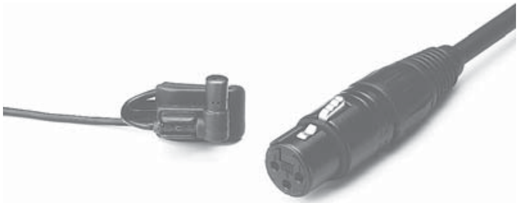


Figure 16-75. Sennheiser MKE 104 lavalier clip-on directional microphone. Courtesy Sennheiser Electronic Corporation.

One of the smallest microphones is the Countryman B6, Fig. 16-77. The B6 microphone has a diameter of 0.1 inches and has replaceable protective caps. Because of its small size, it can be hidden even when it's in plain sight. By choosing a color cap to match the environment, the microphone can be pushed through a button hole or placed in the hair.

Lavalier microphones are normally used to give the talker freedom of movement. This causes problems associated with motion—for instance, noise being transmitted through the microphone cable. To reduce this noise, soft, flexible microphone cable with good fill to reduce wire movement should be used (see Chapter 14). The cable, or power supply for electret/condenser microphones, should be clipped to the user's belt or pants to reduce cable noise to only that created between the clip and the microphone, about 2 ft (0.6 m). Clip-

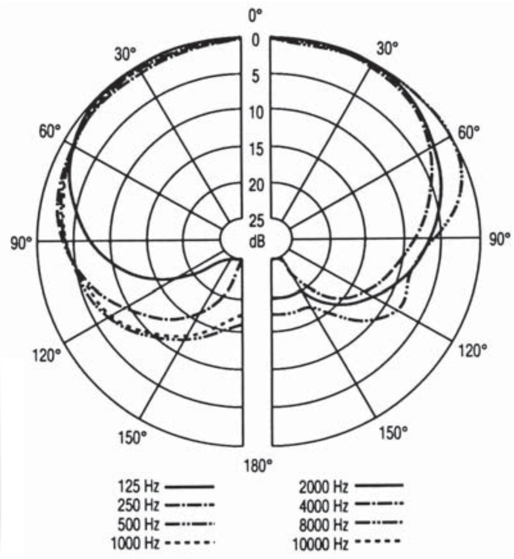


Figure 16-76. Polar response of the microphone in Fig. 16-75. Courtesy Sennheiser Electronic Corporation.

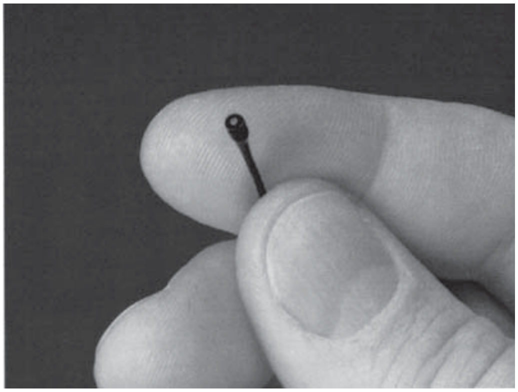


Figure 16-77. Countryman B6 miniature lavalier microphone. Courtesy of Countryman Associates, Inc.

ping to the waist also has the advantage of acting as a strain relief when the cord is accidentally pulled or stepped on.

A second important characteristic of the microphone cable is size. The cable should be as small as possible to make it unobtrusive and light enough so it will not pull on the microphone and clothing.

Because the microphone is normally 10 inches (25 cm) from the mouth of the talker and out of the signal path, the microphone output is less than a microphone on a stand in front of the talker. Unless the torso is between the microphone and loudspeaker, the lavalier microphone is often a prime candidate for feedback. For this reason, the microphone response should be as smooth as possible.

As in any microphone situation, the farther the microphone is away from the source, the more freedom of movement between microphone and source without adverse effects. If the microphone is worn close to the neck for increased gain, the output level will be greatly affected by the raising and lowering and turning of the talker's head. It is important that the microphone be worn chest high and free from clothing, etc. that might cover the capsule, reducing high-frequency response.

16.6.3 Head-Worn Microphones

Head-worn microphones such as the Shure Model SM10A, Fig. 16-78, and Shure Beta 53, Fig. 16-79, are low-impedance, unidirectional, dynamic microphones, designed for sports and news announcing, for interviewing and intercommunications systems, and for special-event remote broadcasting. The Shure SM10A is a unidirectional microphone while the Beta 53 is an omnidirectional microphone.

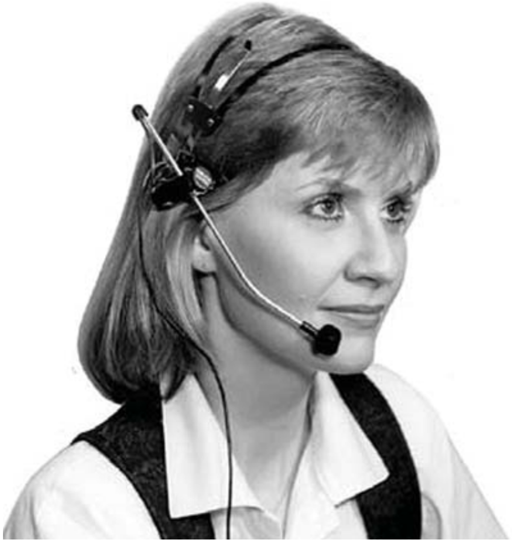


Figure 16-78. Shure SM10A dynamic unidirectional head-worn microphone. Courtesy Shure Incorporated.

Head-worn microphones offer convenient, hands-free operation without user fatigue. As close-talking units, they may be used under noisy conditions without losing or masking voice signals. They are small, lightweight, rugged, and reliable units that

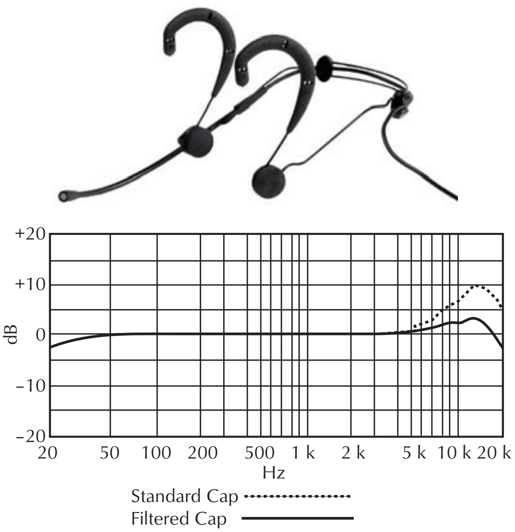


Figure 16-79. Shure Beta 53 omnidirectional microphone and its frequency response. Courtesy Shure Incorporated.

normally mount to a cushioned headband. A pivot permits the microphone boom to be moved 20° in any direction and the distance between the microphone and pivot to be changed 9 cm (3½ in).

Another head-worn microphone, the Countryman Isomax E6 Directional EarSet microphone, is extremely small. The microphone clips around the ear rather than around the head. The units come in different colors to blend in with the background. The ultra-miniature condenser element is held close to the mouth by a thin boom and comfortable ear clip. The entire assembly weighs less than one-tenth of an ounce and almost disappears against the skin, so performers can forget it's there and audiences barely see it, Fig.16-80.

The microphone requires changeable end caps that create a cardioid pickup pattern for ease of placement, or a hypercardioid pattern when more isolation is needed. The C (cardioid) and H (hypercardioid) end caps modify the EarSet's directionality, Fig. 16-81.

The EarSet series should always have a protective cap in place to keep sweat, makeup and other foreign material out of the microphone.

The hypercardioid cap provides the best isolation from all directions, with a null toward the floor where wedge monitors are often placed. The hypercardioid is slightly more sensitive to air movement and handling noise and should always be used with a windscreen.

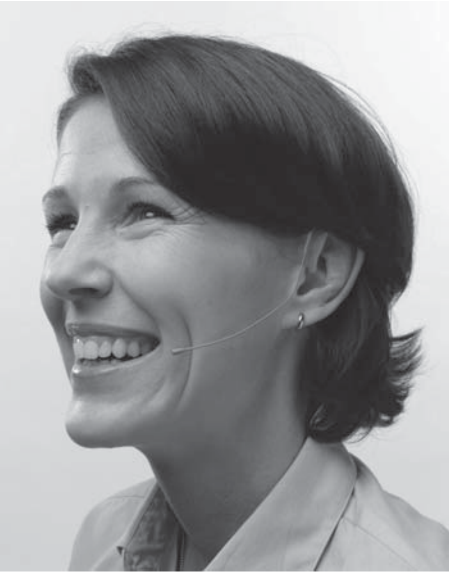
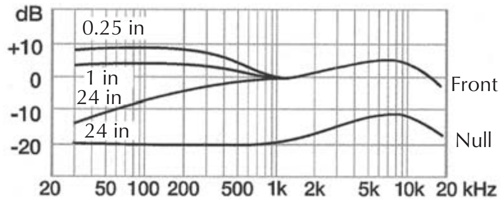
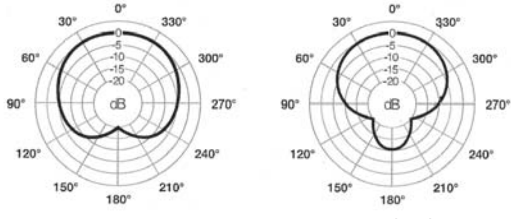


Figure 16-80. Countryman Isomax E6 Directional EarSet microphone. Courtesy Countryman Associates, Inc.



A. Frequency response.



Cardioid cap Hypercardioid cap

B. 1 kHz polar response.

Figure 16-81. Frequency response and polar response of the Countryman E6 EarSet microphone in Fig. 16-80. Courtesy Countryman Associates, Inc.

The cardioid cap is slightly less directional, with a null roughly toward the performer's back. It's useful for trade show presenters or others who have a monitor loudspeaker over their shoulders or behind them.

The microphone can be connected to a sound board or wireless microphone transmitter with the standard 2 mm cable or an extra small 1 mm cable, Fig. 16-80.

16.6.4 Base Station Microphones

Base station power microphones are designed specifically for citizens band transceivers, amateur radio, and two-way radio applications. For clearer transmission and improved reliability, transistorized microphones can be used to replace ceramic or dynamic, high- or low-impedance microphones supplied as original equipment.

The Shure Model 450 Series II, Fig. 16-82, is a high output dynamic microphone designed for paging and dispatching applications. The microphone has an omnidirectional pickup pattern and a frequency response tailored for optimum speech intelligibility, Fig. 16-83. It includes an output impedance selection switch for high, 30,000 Ω, and low, 225 Ω, and a locking press-to-talk switch.

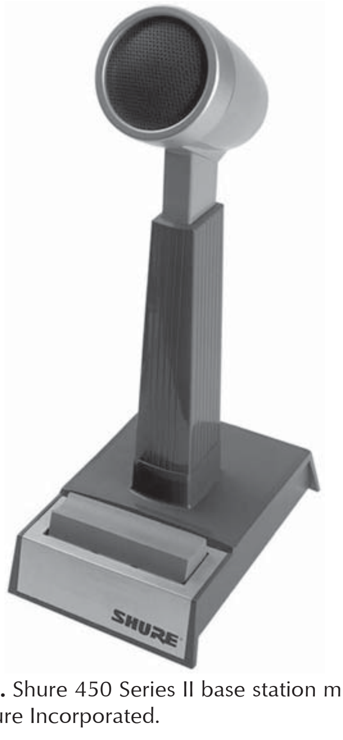


Figure 16-82. Shure 450 Series II base station microphone. Courtesy Shure Incorporated.

The press-to-talk switch can be converted to a monitor/transmit switch with the Shure RK199S Split-Bar Conversion Kit. When the optional split-bar Transmit/Monitor Switch Conversion Kit is installed, the monitor bar must be depressed before the transmit switch can be depressed, requiring the operator to verify that the channel is open before transmitting. The monitor bar can be locked in the on position. The transmit bar is momentary and cannot be locked.

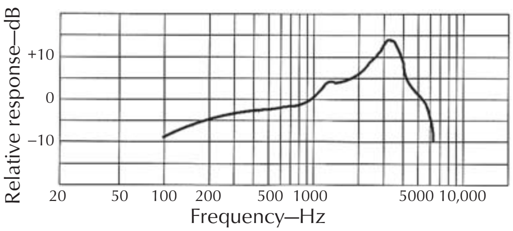


Figure 16-83. Frequency response of the Shure 450 Series II microphone shown in Fig. 16-82. Courtesy Shure Incorporated

16.6.5 Differential Noise-Canceling Microphones

Differential noise-canceling microphones, Fig. 16-84, are essentially designed for use in automobiles, aircraft, boats, tanks, public-address systems, industrial plants, or any service where the ambient noise level is 80 dB or greater and the microphone is handheld. Discrimination is afforded against all sounds originating more than ¼ in (6.4 mm) from the front of the microphone. The noise-canceling characteristic is achieved through the use of a balanced port opening, which directs the unwanted sound to the rear of the dynamic unit diaphragm out of phase with the sound arriving at the front of the microphone. The noise canceling is most effective for frequencies above 2000 Hz. Only speech originating within ¼ in (6.4 mm) of the aperture is fully reproduced. The average discrimination between speech and noise is 20 dB with a frequency response of 200–5000 Hz.

16.6.6 Controlled-Reluctance Microphones

The *controlled-reluctance microphone* operates on the principle that an electrical current is induced in a coil, located in a changing magnetic field. A magnetic armature is attached to a diaphragm suspended inside a coil. The diaphragm, when disturbed by a sound wave, moves the armature and induces a corresponding varying voltage in the coil. High output with fairly good frequency response is typical of this type of microphone.

16.6.7 Handheld Entertainer Microphones

The *handheld entertainer microphone* is most often used by a performer on stage and, therefore, requires a special frequency response that will increase articulation and presence. The microphones are often subjected to rough handling, extreme shock, and vibration. For live performances, the proximity effect can be useful to produce a low bass sound.

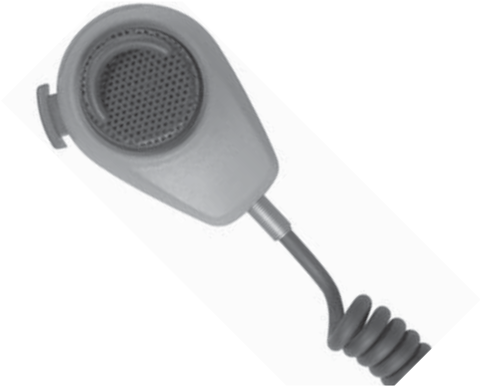


Figure 16-84. Shure 577B dynamic noise-canceling microphone. Courtesy Shure Incorporated.

Probably the most famous entertainer's microphone is the Shure SM58, Fig. 16-85. The microphone has a highly effective spherical wind screen that also reduces breath pop noise. The cardioid pickup pattern helps reduce feedback. The frequency response, Fig. 16-86, is tailored for vocals with brightened midrange and bass roll-off. Table 16-2 gives the suggested microphone placement for best tone quality.

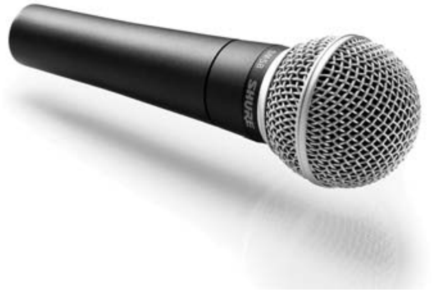


Figure 16-85. Shure SM58 vocal microphone. Courtesy Shure Incorporated.

To overcome rough handling and handling noise, special construction techniques are used to reduce wind, pop noise, and mechanical noise and to ensure that the microphone will withstand sudden collisions with the floor. The Sennheiser MD431, Fig. 16-87, is an example of a high-quality, rugged, and low-mechanical-noise microphone. To eliminate feedback, the MD431 incorporates a supercardioid directional characteristic, reducing side pickup to 12% or less than half that of conventional cardioids.

Another problem, particularly with powerful sound reinforcement systems, is mechanical (handling) noise. Aside from disturbing the audience, it can actually damage equipment. As can be seen in the cutaway, the MD 431 is actually a microphone within a microphone. The dynamic transducer element is mounted within an inner capsule, isolated from the outer housing by means of a shock absorber. This protects it from handling noise as well as other mechanical vibrations normally encountered in live performances.

Table 16-2. Suggested Placement for the SM58 Microphone

Application	Suggested Microphone Placement	Tone Quality
Lead and backup vocals	Lips less than 150 mm (6 in) away or touching the windscreen, on axis to microphone	Robust sound, emphasized bass, maximum isolation from other sources
Speech	from mouth, just above nose height	Natural sound, reduced bass
	200 mm (8 in) to 0.6 m (2 ft) away from mouth, slightly off to one side	Natural sound, reduced bass and minimal "s" sounds
	1 m (3 ft) to 2 m (6 ft) away	Thinner; distant sound; ambience

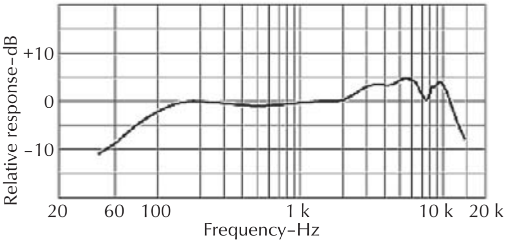


Figure 16-86. Frequency response of the Shure SM58 vocal microphone. Courtesy Shure Incorporated.

To screen out noise still further, an internal electrical high pass filter network is incorporated to insure that low-frequency disturbances will not affect the audio signal. A built-in mesh filter in front of the diaphragm reduces the popping and excessive sibilance often produced by close micing.

The microphone case is a heavy-duty cast outer housing with a stainless steel front grille and reed-type on–off switch. A hum bucking coil is mounted behind the transducer to cancel out any stray magnetic fields.

16.6.8 Pressure-Gradient Condenser Microphones

One of the most popular studio microphones is the Neumann U-87 multidirectional condenser microphone, Fig. 16-88, and its cousin, the Neumann U-89,

Fig. 16-89. This microphone is used for close micing where high SPLs are commonly encountered. The response below 30 Hz is rolled off to prevent low-frequency blocking and can be switched to 200 Hz to allow compensation for the bass rise common to all directional microphones at close range.

The figure-eight characteristic is produced by means of two closely spaced or assembled cardioid characteristic capsules, whose principal axes are pointed in opposite directions and are electrically connected in antiphase.

These microphones are usually made with backplates equipped with holes, slots, and chambers forming delay elements whose perforations act as part friction resistances and part energy storage (acoustic inductances and capacitances), giving the backplate the character of an acoustic low-pass network. In the cutoff range of this low-pass network, above the transition frequency f_t, the membrane is impinged upon only from the front, and the microphone capsule changes to a pressure or interference transducer.

The output voltage $e(t)$ of a condenser microphone using dc polarization is proportional to the applied dc voltage E_o and, for small displacement amplitudes of the diaphragm, to the relative variation in capacity $c(t)/C_o$ caused by the sound pressure

$$\mathrm{e(t)} = E_o\frac{c(t)}{C_o} \tag{16-25}$$

where,

E_o is the applied dc voltage,

$c(t)$ is the variable component of capsule capacity,

C_o is the capsule capacity in the absence of sound pressure,

t is the time.

The dependence of output voltage $e(t)$ on E_o is utilized in some types of microphones to control the directional characteristic. Two capsules with cardioid characteristics as shown in Fig. 16-90 are placed back to back. They can also be assembled as a unit with a common backplate. The audio (ac) signals provided by the two diaphragms are connected in parallel through a capacitor C. The intensity and phase relationship of the outputs from the two capsule halves can be affected by varying the dc voltage applied to one of them (the left cartridge in Fig. 16-90). This can be accomplished through a switch, or a potentiometer. The directional characteristic of the microphone may be changed by remote control via long extension cables.

If the switch is in its center position C, then the left capsule-half does not contribute any voltage, and the microphone has the cardioid characteristic of the right capsule-half. In switch position A, the two ac voltages are in parallel, resulting in an omnidirectional pattern. In position E the two halves are connected in antiphase, and the result is a figure-8 directional response pattern.

The letters A to E given for the switch positions in Fig. 16-90 produce the patterns given the same letters in Fig. 16-91.

16.6.9 Interference Tube Microphone

The *interference tube microphone*[8] as described by Olson in 1938 is often called a *shotgun microphone* because of its physical shape and directional characteristics.

Important characteristics of any microphone are its sensitivity and directional qualities. Assuming a constant sound pressure source, increasing the micro-

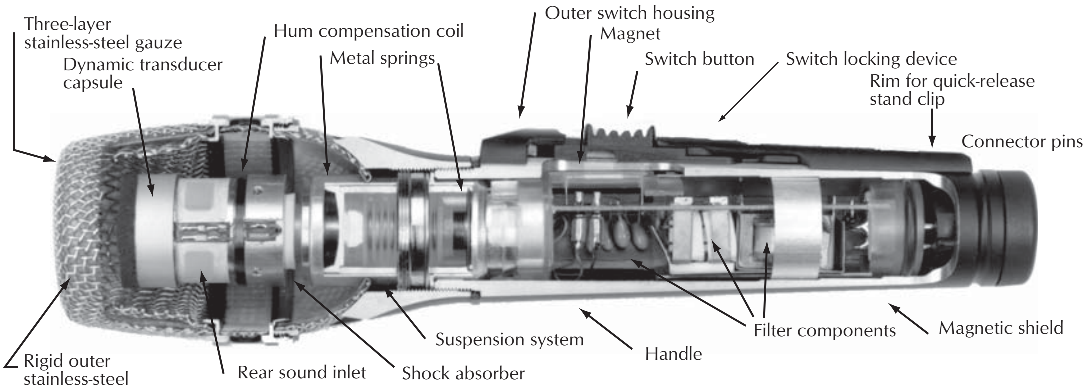


Figure 16-87. Cut-away view of a Sennheiser MD 431 handheld entertainment microphone. Courtesy Sennheiser Electronic Corporation.

NEUMANN-MICROPHON
4905

Figure 16-88. Neumann U-87 microphone. Courtesy Neumann USA.

phone to the source distance requires an increase in the gain of the amplifying system to produce the same output level. This is accompanied by a decrease in SNR and an increase in environmental noises including reverberation and background noise to where the indirect sound may equal the direct sound. The wanted signal then deteriorates to where it is unusable. Distance limitations can be overcome by increasing the sensitivity of the microphone, and the effect of reverberation and noise pickup can be lessened by increasing the directivity of the pattern. The interference tube microphone has these two desirable qualities.

The DPA 4017 is a supercardioid shotgun microphone. It is 8.3 in (210 mm) long and weighs 2.5 oz (71 g), making it useful on booms, Fig. 16-92. The polar pattern is shown in Fig. 16-93.

The difference between interference tube microphones and standard microphones lies in the method of pickup.

An interference tube is mounted over the diaphragm and is schematically drawn in Fig. 16-94.

The microphone consists of four parts as shown in the schematic:

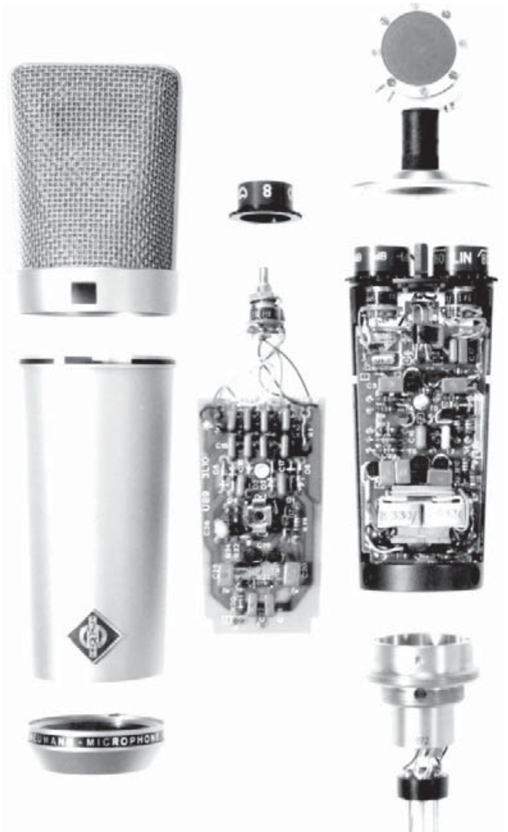


Figure 16-89. Neumann U-89 microphone. Courtesy Neumann USA.

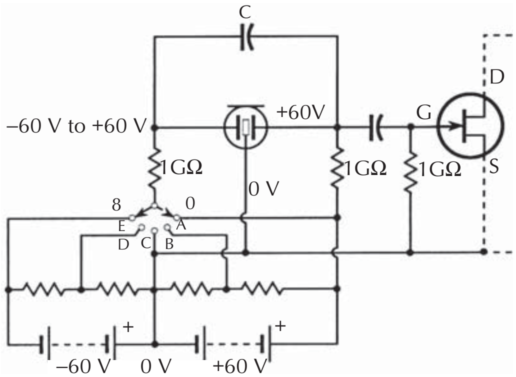


Figure 16-90. Circuit of the Neumann U-87 condenser microphone with electrically switchable direction characteristic. Courtesy Neumann USA.

1. Interference tube with one frontal and several lateral inlets covered with fabric or other damping material.

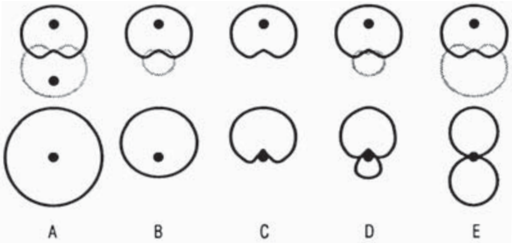
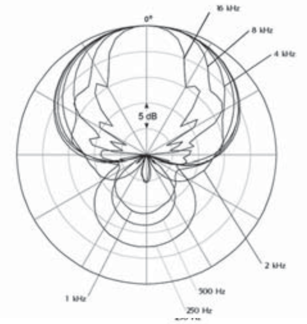
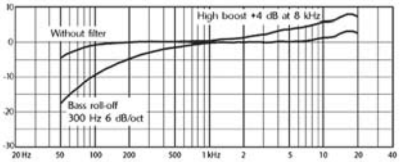


Figure 16-91. By using a microphone as shown in Fig. 16-89 and superimposing two cardioid patterns (top row), directional response patterns (bottom row) can be obtained. Courtesy Neumann USA.

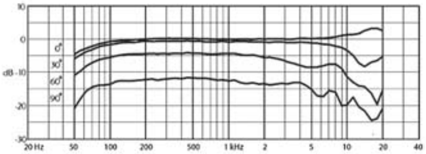
Figure 16-92. A supercardioid interference tube microphone. Courtesy DPA Microphones, Inc.



A. Directional characteristics.



B. Frequency response of switching filters.



C. On and off axis response measured at 60 cm (23.6 in).

Figure 16-93. Directional characteristics and frequency response of a DPA 4017 microphone. Courtesy DPA Microphones, Inc.

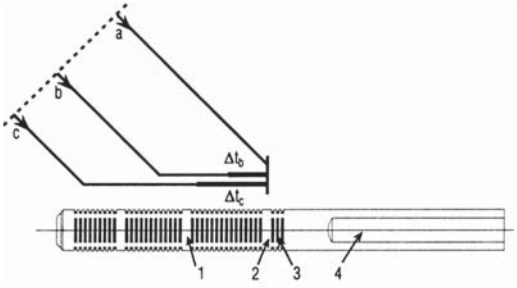


Figure 16-94. Schematic of an interference tube microphone.

2. Capsule with the diaphragm and counter electrode(s).
3. Rear inlet.
4. Electronic circuit.

The directional characteristics are based on two different principles:

1. In the low-frequency range, tube microphones behave as first-order directional receivers. The tube in front of the capsule can be considered as an acoustic element with a compliance due to the enclosed air volume and a resistance determined by the lateral holes or slits of the tube. The rear inlet is designed as an acoustic low-pass filter to achieve the phase shift for the desired polar pattern (normally cardioid or supercardioid).
2. In the high-frequency range, the acoustical properties of the interference tube determine the polar patterns. The transition frequency between the two different directional characteristics depends on the length of the tube and is given by

$$f_o = \frac{c}{2L} \tag{16-26}$$

where,

f_o is the transition frequency,

c is the velocity of sound in air in feet per second or meters per second,

L is the length of the interference tube in feet or meters.

Referring to Fig. 16-94, if the tube is exposed to a planar sound wave, every lateral inlet is the starting point of a new wave traveling inside the tube toward the capsule as well as towards the frontal inlet. Apart from the case of frontal sound incidence, every particular wave covers a different distance to the capsule and, therefore, arrives at a different time. Fig. 16-94 shows the delay times of waves *b* and *c* compared to wave *a*. Note that they increase with the angle of sound incidence.

The resulting pressure at the capsule can be calculated by the sum of all particular waves generated over the tube's length, all with equal amplitudes but different phase shifts. The frequency and phase response curves can be described by

$$\frac{P\theta}{P(\theta = 0°)} = \frac{\sin\left[\frac{\pi L}{\lambda} \times (1 - \cos\theta)\right]}{\frac{\pi L}{\lambda} \times (1 - \cos\theta)} \qquad (16\text{-}27)$$

where,

$P(\theta)$ is the microphone output at a given angle of sound incidence,
$P(\theta = 0°)$ is the microphone output along principal axis,
λ is the wavelength,
L is the length of the tube,
θ is the angle of sound incidence.

The calculated curves and polar patterns are plotted in Figs. 16-95 and 16-96 for a tube length of 9.8 inches (25 cm) without regard to the low-frequency directivity caused by the rear inlet. The shape of the response curves looks similar to that of a comb filter with equidistant minima and maxima decreasing with 6 dB/octave. The phase response is frequency independent only for frontal sound incidence. For other incidence angles, the phase depends linearly on frequency, so that the resulting pressure at the capsule shows an increasing delay time with an increasing incidence angle.

In practice, interference tube microphones show deviations from this simplified theoretical model. Fig. 16-97 is the polar pattern of the Sennheiser MKH 60P48. The built-in tube delivers a high-frequency roll-off for lateral sound incidence with a sufficient attenuation especially for the first side lobes. The shape of the lateral inlets as well as the covering material influences the frequency and phase response curves. The transition frequency can be lowered with an acoustic mass in the frontal inlet of the tube to increase the delay times for low frequencies.

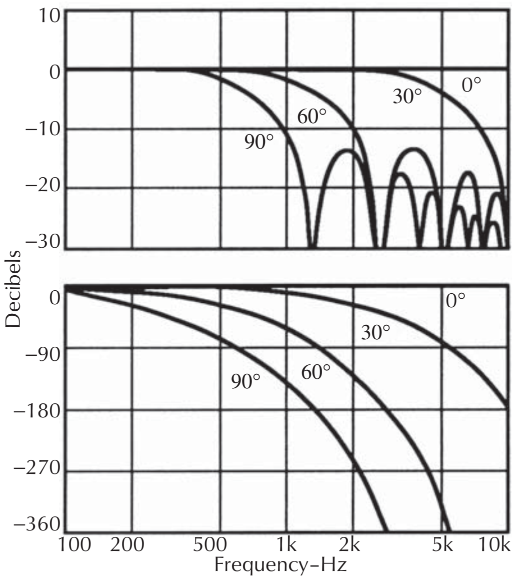


Figure 16-95. Calculated frequency and phase response curves of an interference tube microphone (250 mm) without rear inlet for different angles of sound incidence. Courtesy Sennheiser Electronic Corporation.

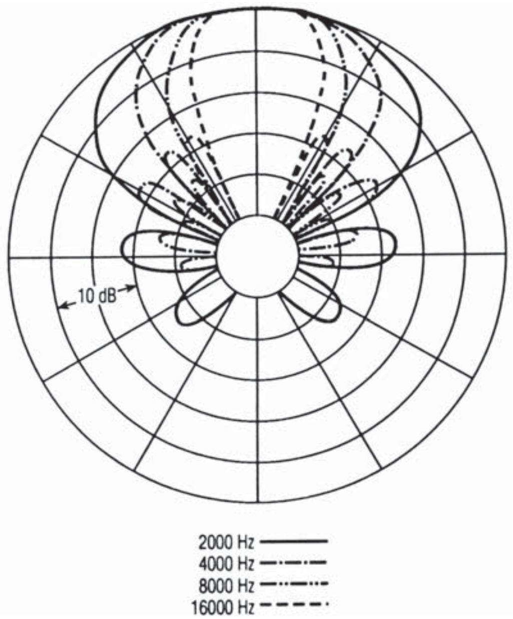


Figure 16-96. Calculated polar patterns of an interference tube microphone (250 mm) without rear inlet. Courtesy Sennheiser Electronic Corporation.

Another interference tube microphone is the Shure SM89, Fig. 16-98.[9] In this microphone, a tapered acoustic resistance is placed over the elongated interference tube slit, varying the effective length of the tube with frequency so that *L*/*M* (the ratio of tube length to wavelength) remains nearly constant over the desirable frequency range. This allows the polar response to be more consistent as frequency increases, Fig. 16-99, because the resistance in conjunction with the compliance of the air inside the tube forms an acoustical low-pass filter. High frequencies are attenuated at the end of the tube because it is the high-resistance end,

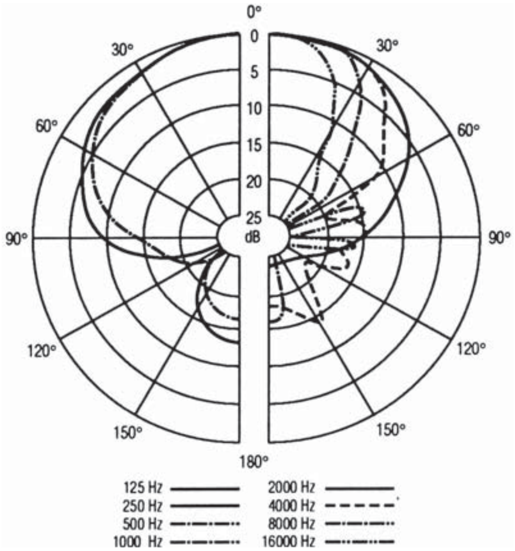


Figure 16-97. Characteristic of a supercardioid microphone (MKH 60). Courtesy Sennheiser Electronic Corporation.

Figure 16-98. Shure SM89 condenser shotgun microphone. Courtesy Shure Incorporated.

allowing the high frequencies to enter the tube only near the diaphragm. This makes the tube look shorter at high frequencies, Eq. 16-27.

While a cardioid microphone may be capable of picking up satisfactorily at 3 ft (1 m), a cardioid in-line may reach 6–9 ft (1.8–2.7 m), and a super in-line may reach as far as 40 ft (12 m) and be used for picking up a group of persons in a crowd from the roof of a nearby building, following a horse around a race track, picking up a band in a parade, and picking up other hard-to-get sounds from a distance.

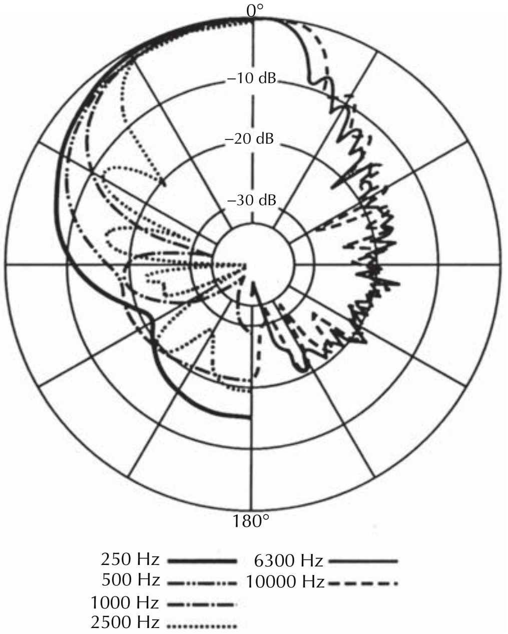


Figure 16-99. Polar response of a 0.5 m long shotgun microphone. Courtesy Shure Incorporated.

There are precautions that should be followed when using interference tube microphones. Because they obtain directivity by cancellation, frequency response and phase are not as smooth as omnidirectional microphones. Also, since low frequencies become omnidirectional, the frequency response drops rapidly below 200 Hz, which helps control directivity.

It should not be assumed that no sound will be picked up outside the pickup cone. As the microphone is rotated from an on-axis position to a 180° off-axis position, there will be a progressive drop in level. Sounds originating at angles of 90° to 180° off-axis will cancel by 20 dB or more; however, the amount of cancellation depends on the level and distance of the microphone from the sound source. As an example, if an on-axis sound originated at a distance of 20 ft (6 m), a 90° to 180° off-axis sound occurring at the same distance and intensity will be reduced by 20 dB or more, providing none of the off-axis sound is reflected into the front of the microphone by walls, ceiling, and so on. On the other hand, should the off-axis sound originate at a distance of 2 ft (0.6 m) and at the same sound pressure level as the sound at 20 ft (6 m) on axis, it will be reproduced at the same level. The reason for this behavior is that the microphone is still canceling the unwanted sound as much as 20 dB, but due to the difference in the distances of the two sounds, the off-axis sound is 20 dB louder than the on-axis sound at the microphone. Therefore, they are reproduced at the same level. For a pickup in an area where random noise and reverberation are problems, the microphone should be located with the back end to the source of unwanted sound and as far from the disturbances as possible.

If the microphone is being used inside a truck or enclosed area, and pointing out a rear door, poor pickup may be experienced because all sounds, both wanted and unwanted, arrive at the microphone on-axis. Since the only entrance is through the truck door, no cancellation occurs because the truck walls inhibit the sound from entering the sides of the microphone. In this instance, the microphone will be operating as an omnidirectional microphone. Due to the reflected sound from the walls, the same condition will prevail in a room where the microphone is pointed through a window or when operating in a long hallway. For good pickup, the microphone should be operated in the open and not in closely confined quarters.

A shotgun interference tube microphone cannot be compared to a zoom lens since the focus does not vary nor does it reach out to gather in the sound. What the narrow polar pattern and high rate of cancellation do are to reduce pickup of the random sound energy and permit the raising of the amplifier gain following the microphone without seriously decreasing the SNR.

Difficulties may also be encountered using interference tube microphones on stage and picking out a talker in the audience, particularly where the voice is 75–100 ft (23–30 m) away and fed back through a reinforcement system for the audience to hear. Under these circumstances, only about 30–50 ft (9–15 m) is possible without acoustic feedback; even then, the system must be balanced very carefully.

16.6.10 Rifle Microphones

The *rifle microphone* consists of a series of tubes of varied length mounted in front of either a capacitor or dynamic transducer diaphragm, Fig. 16-100. The transducer may be either a capacitor or dynamic type. The tubes are cut in lengths from 2–60 inches (5–150 m) and bound together. The bundling of the tubes in front of the transducer diaphragm creates a distributed sound entrance, and the omnidirectional transducer becomes highly directional.

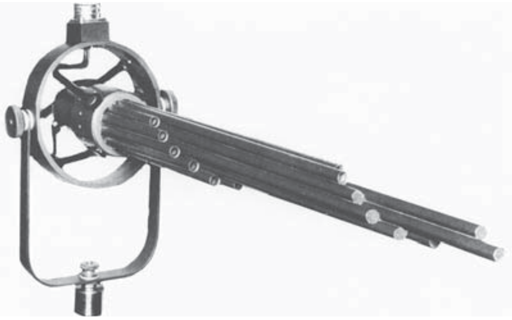


Figure 16-100. RCA rifle microphone. Courtesy of Radio Corporation of America.

Sound originating on the axis of the tubes first enters the longest tube and, as the wavefront advances, enters successively shorter tubes in normal progression until the diaphragm is reached. Sounds reaching the diaphragm from the source travel the same distance, regardless of the tube entered, so all sounds arriving on-axis are in phase when they reach the diaphragm. Sounds originating 90° off-axis enter all tubes simultaneously. A sound entering a longer tube may travel 18 inches (46 cm) to reach the diaphragm, while the same sound traveling through the shortest tube will travel only 3 inches (7.6 cm), with other differences for the varied length of tubing causing an out-of-phase signal at the diaphragm. Under these conditions, a large portion of the sound originating at 90° is canceled, and from 180° an even greater phase difference occurs, and cancellation is increased considerably.

The RCA MI-100006A varidirectional microphone, Fig. 16-100, consists of nineteen $^5/_{16}$ inches (0.8 cm) plastic tubes, ranging from 3–18 inches (7.6–46 cm) in length. The tubes are bundled and mounted in front of an omnidirectional capacitor-microphone head. Rifle microphones are not used very much today.

16.6.11 Parabolic Microphones

Parabolic microphones use a parabolic reflector with a microphone to obtain a highly directional pickup response. The microphone diaphragm is mounted at the focal point of the reflector, Fig. 16-101. Any sound arriving from an angle other than straight on will be scattered and therefore will not focus on the pickup. The microphone is focused by moving the diaphragm in or out from the reflector for maximum pickup. This type concentrator is often used to pick up a horse race or a group of people in a crowd.

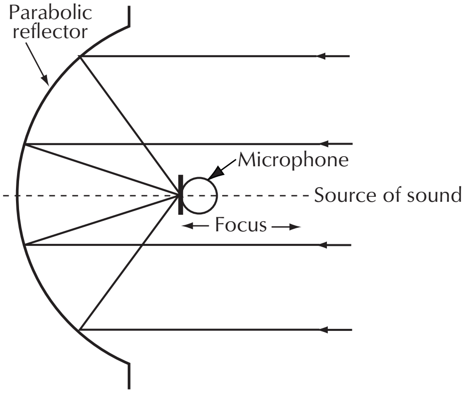


Figure 16-101. A parabolic bowl concentrator for directional microphone pickup.

The greatest gain in sound pressure is obtained when the reflector is large compared to the wavelength of the incident sound. With the microphone in focus, the gain is the greatest at the mid-frequency range. The loss of high frequencies may be improved somewhat by defocusing the microphone a slight amount, which also tends to broaden the sharp directional characteristics at the higher frequencies. A bowl 3 ft (0.91 m) in diameter is practically nondirectional below 200 Hz but is very sharp at 8000 Hz, Fig. 16-102. For a diameter of 3 ft, the gain over the microphone without the bowl is about 10 dB and, for a 6 ft (1.8 m) diameter bowl, approximately 16 dB.

16.6.12 Zoom Microphones

A *zoom microphone*,[10] or *variable-directivity microphone*, is one that operates like and in conjunction with a zoom lens. This type of microphone is useful with television and motion-picture operations.

The optical perception of distance to the object is simply determined by the shot angle of the picture. On the other hand, a sound image is perceived by:

- Loudness.
- Reverberation (ratio of direct sound to reflected sound).
- Acquired response to sound.

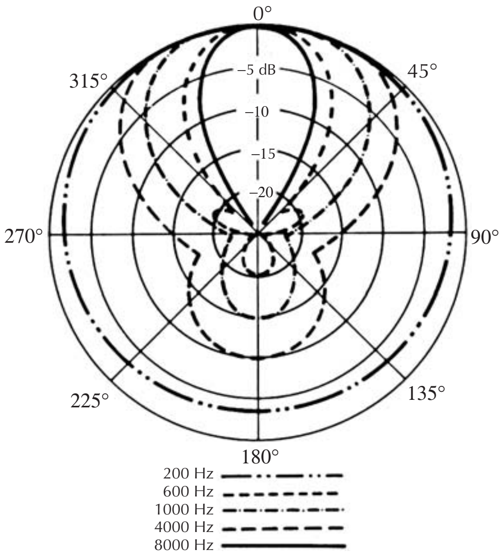


Figure 16-102. Polar pattern for a parabolic concentrator.

- Level and arriving time difference between the two ears.

If the sound is recorded in monophonic, the following factors can be skillfully combined to reproduce a natural sound image with respect to the perceived distance:

- *Loudness*: Perceived loudness can be controlled by varying microphone sensitivity.
- *Reverberation*: The representation of the distance is made by changing the microphone directivity or the ratio between direct and reverberant sound. In a normal environment, we hear a combination of direct sound and its reflections. The nearer a listening point is to the source, the larger the ratio of direct to reverberant sound. The farther the listening point is from the source, the smaller the ratio; therefore, use of a high-directivity microphone to keep direct sound greater than reflected sound permits the microphone to get apparently closer to the source by reducing reverberant sound pickup. For outdoor environments, use of directional microphones allows the ambient noise level to be changed for natural representation of distances.
- *Acquired human response to sound*: Normally we can tell approximately how far a familiar object as a car or a person is by the sound generated by the objects because we acquire the response to sound through our daily experiences.

16.6.12.1 Distance Factor

The fact that microphone directivity determines the perceived distance can be explained from the viewpoint of the distance factor. Fig. 16-103 shows the sound pressure level at the position of an omnidirectional microphone versus the distance between the microphone and a sound source *S*, with ambient noise evenly distributed. Suppose the distance is 23 ft (7 m) and the ambient noise level is at 1. If the microphone is replaced by one that has a narrow directivity with the same on-axis sensitivity, less noise is picked up, so the observed noise level is lowered to 2. For an omnidirectional microphone, the same effect can be obtained at a distance of 7.5 ft (2.3 m). From a different standpoint, the same SNR as for an omnidirectional microphone at 20.6 ft (6.3 m) can be obtained at a distance of 65 ft (20 m). The ratio of actual-to-observed distance is called the *distance factor*.

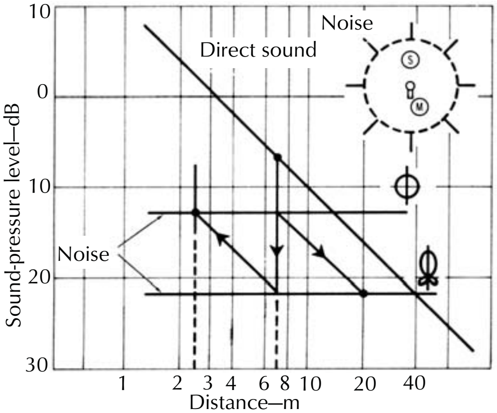


Figure 16-103. Relationship between sound pressure level and distance in an evenly distributed noise environment.

16.6.12.2 Operation of Zoom Microphones

By changing the sensitivity and directivity of a microphone simultaneously, an acoustical zoom effect is realized, and more reality becomes possible in sound recording. Fig. 16-104 is the basic block diagram of a zoom microphone system. The system consists of three unidirectional microphone capsules (1 through 3) arranged on the same axis. The three capsules have the same characteristics, and capsule 3 faces the opposite direction. The directivity can be varied from omnidirectional to second-order gradient unidirectional by varying the mixing ratio of the output of each capsule and changing the equalization characteristic accordingly. An omnidirectional pattern is obtained by simply combining the outputs of capsule 2 and 3. In the process of directivity change from omnidirectional to unidirectional, the output of capsule 3 is gradually faded out, while the output of capsule 1 is kept off. Furthermore, the equalization characteristic is kept flat, because the on-axis frequency response does not change during this process. In the process of changing from unidirectional to second-order gradient unidirectional, the output of capsule 3 is kept off. The second-order gradient unidirectional pattern is obtained by subtracting the output of capsule 1 from the output of capsule 2. To obtain the second-order gradient unidirectional pattern with minimum error, the output level of capsule 1 needs to be trimmed. Since the on-axis response varies according to the mixing ratio, the equalization characteristics also have to be adjusted along with the level adjustment of the output of capsule 1. The on-axis sensitivity increase of second-order gradient setup over the unidirectional setup allows the gain of the amplification to be unchanged.

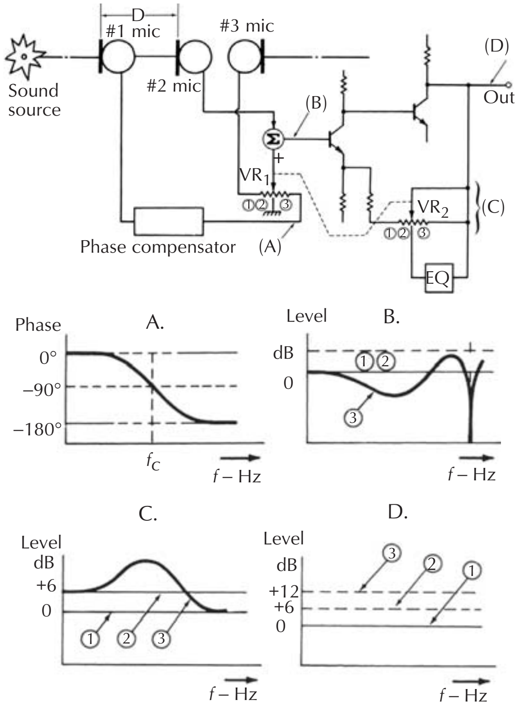


Figure 16-104. Configuration of the zoom microphone.

16.6.12.3 Zoom Microphone Video Camera Linkage

In order to obtain a good matching of picture and sound, a mechanism that synchronizes the optical zooming and acoustical zooming becomes inevitable. Electrical synchronization would also be possible by using voltage-controlled amplifiers (VCA) or voltage-controlled resistors (VCR).

16.6.13 Automatic Microphone Systems

There have been many advances in automatic mixers where the microphone is normally off until gated on by a signal, hopefully, a wanted signal. Many operate on an increased level in one or more microphones with respect to the random background noise, see Chapter 21.

While the Shure *Automatic Microphone System (*AMS*)* is a discontinued microphone, it is still used in many venues. The system turns microphones on and off (with automatic gating), greatly reducing the reverberant sound quality and feedback problems often associated with the use of multiple microphones. The AMS microphones are gated on only by sounds arriving from the front within their acceptance angle of 120°. Other sounds outside the 120° angle, including background noise, will not gate the microphones on, regardless of level. In addition, the AMS adjusts gain automatically to prevent feedback as the number of on microphones increases.

The Shure Model AMS22 low-profile condenser microphone, Fig. 16-105, is designed for use only with the Shure AMS. Unlike conventional microphones, it contains electronic circuitry and a novel transducer configuration to make it compatible with the Shure AMS mixers. The microphone should not be connected to standard simplex- (phantom-) or non-simplex-powered microphone inputs because they will not function properly.

AMS microphones, in conjunction with the special circuitry in the AMS mixers, uniquely discriminate between desired sounds that originate within their 120° front acceptance angle and all other sounds. Sounds from the front of a microphone are detected and cause it to be gated on, transmitting its signal to the mixer output. Sounds outside the acceptance angle will not gate the microphone on. When an AMS22 is gated on, it operates like a hemi- or half-cardioid microphone because half the cardioid pattern disappears when the microphone is placed on a surface, Fig. 16-106. Each AMS microphone operates completely independently in analyzing its own sound field and deciding whether or not a sound source is within the front acceptance angle.

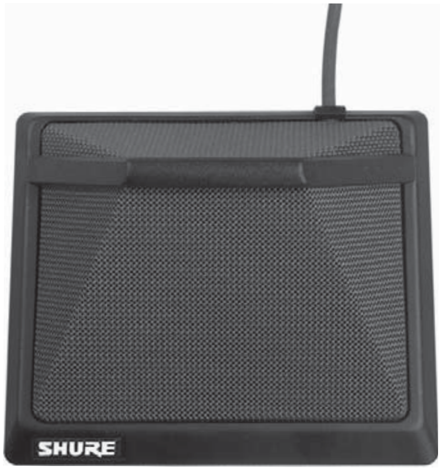


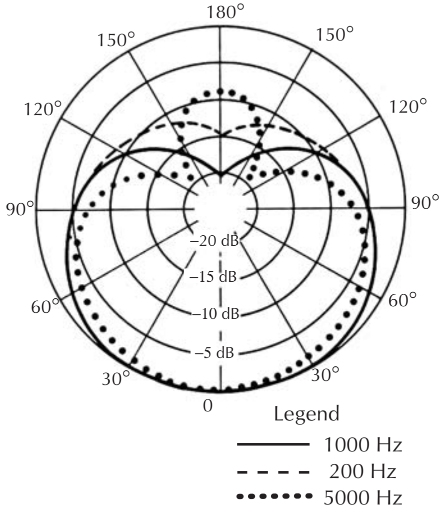
Figure 16-105. Shure Automatic Microphone System (AMS) model AMS22 low-profile microphone. Courtesy Shure Incorporated.

The microphone should be placed so that intended sources are within 60° of either side of the front of the microphone—that is, within 120° acceptance angle. Sources of undesired sound should be located outside the 120° acceptance angle. Each microphone should be at least 3 ft from the wall behind it, and items such as large ashtrays or briefcases should be at least 1 ft behind it. If the microphones are closer than that, reflections will reduce the front-to-back discrimination and, therefore, make the microphone act more like a conventional cardioid type.

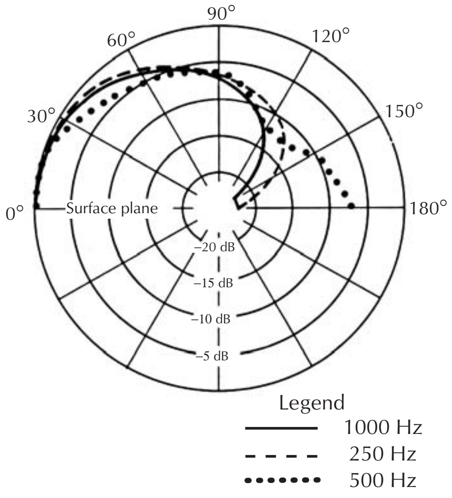
16.6.14 PolarFlex™ Microphone System

The PolarFlex system by Schoeps models any microphone. The system features two output channels with two microphones per channel, Fig. 16-107. The standard system consists of an omnidirectional and a figure 8 microphone for each channel and an analog/digital processor.

Essential sonic differences between condenser microphones of the same nominal directional pattern are not only due to frequency response, but also to the fact that the polar pattern is not always uniformly maintained throughout the entire frequency range particularly at the lowest and highest frequencies. Though ostensibly a defect, this fact can also be used to advantage (e.g., adaptation to the acoustic of the recording room). While the frequency response at a given pickup angle can be controlled by equalizers, there was no way to alter the polar pattern correspondingly. The only way



A. Side view



B. Top view

Figure 16-106. Hemicardioid polar pickup pattern for a Shure AMS surface microphone. Courtesy Shure Incorporated.

to control this situation was through the choice of microphones having different variations of the polar pattern versus frequency. With the DSP-4P processor, nearly ideal directional characteristics can be selected, and nearly any frequency-dependent directional characteristic that may be desired—e.g., a cardioid becomes omnidirectional below the midrange, so that it has better response at the very lowest frequencies. Also modeling a large-diaphragm microphone is possible.

Furthermore, in excessively reverberant spaces one could record a drier sound (cardioid or supercardioid setting) or, in spaces that are dry, accept more room reflections (wide cardioid or omni setting) in the corresponding frequency range.

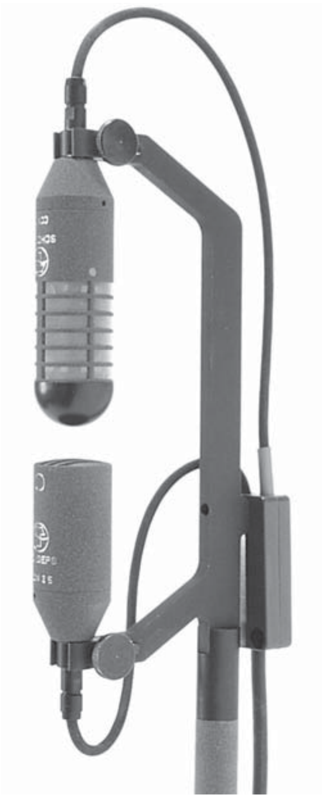


Figure 16-107. A Schoeps PolarFlex microphone with an omnidirectional and a figure 8 microphone. Courtesy Schoeps GmbH.

In such cases it is not the frequency response but rather the ratio of direct to reflected sound, that will be altered. That cannot be done with an equalizer nor can a reverb unit reduce the degree of reflected sound after the fact.

In the arrangement of Fig. 16-107, an omnidirectional microphone with a mild high-frequency emphasis in the direct sound field is used. Because of its angle of orientation, the capsule has ideal directional response in the horizontal plane; the high-frequency emphasis compensates for the high-frequency losses due to lateral sound incidence.

A figure 8 microphone is set directly above the omni. The direction to which it is aimed will determine the orientation of the resulting adjustable virtual microphone. The hemispherical device attached to the top of the figure 8 flattens the response of the omnidirectional microphone at the highest frequencies.

By using the DSP-4P processor, Fig. 16-108, the following settings can be made independently of one another in three adjustable frequency ranges. With the three knobs in the upper row, the directional patterns in each of the three frequency bands can be set. The settings are indicated by a circle of LEDs around each of the knobs. At the lower left of each knob is the omnidirectional setting; at the lower right is the figure 8 setting. Eleven intermediate pattern settings are available. The knobs in the lower row are set between those in the upper row. They are used for setting the boundaries between the frequency ranges; 100 Hz–1 kHz and 1–10 kHz, respectively, in ⅓ octave steps.

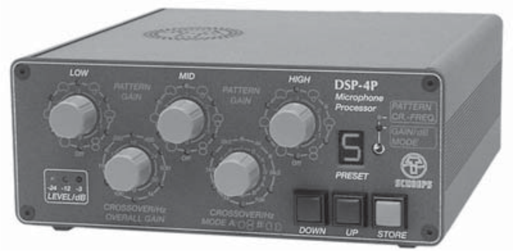


Figure 16-108. Schoeps DSP-4P microphone processor. Courtesy Schoeps GmbH.

The three buttons at the lower right are for storing and recalling presets. If the unprocessed microphone signals have been recorded, these adjustments can be made during postprocessing.

The processor operates at 24-bit resolution with either a 44.1 kHz or 48 kHz sampling rate. When a digital device is connected to the input, the PolarFlex™ processor adapts to its clock signal.

16.7 Stereo Microphones

Stereo microphones are microphones or systems used for coincident, XY, M/S, SASS, binaural in-the-head, and binaural in-the-ear (ITE) recording. These systems have the microphones close together (in proximity of a point source or ear-to-ear distance) and produce the stereophonic effect by intensity stereo, time-based stereo, or a combination of both.

16.7.1 Coincident Microphones

A highly versatile stereo pickup is the *coincident microphone technique*.[11,12,13] Coincident means that sound reaches both microphones at the same time, implying that they are at the same point in space. In practice, the two microphones cannot occupy the same point, but they are placed as closely together as possible. There are special-purpose stereo microphones available that combine the two microphones in one case. Since they are essentially at the same point, there can be no time differences between arrival of any sound from any direction; thus no cancellation can occur. It might first appear that there could be no stereophonic result from this configuration. The two microphones are usually unidirectional and oriented at 90° to one another. The combination is then aimed at the sound source, each microphone 45° to a line through the source. Stereo results from intensity differences—the left microphone (which is to the right of the pair) will receive sounds from the left-hand part of the stage with greater volume than it will receive from the right-hand side of the stage.

The stereo result, although often not as spectacular as that obtained from spaced microphones, is fully mono compatible, and it most accurately reproduces the sound of the acoustic environment. It is quite foolproof and quick to set up.

Variations of the coincident technique include changing the angle between the microphone (some stereo microphones are adjustable); using bidirectional microphones, which results in more reverberant sound; using combinations of unidirectional and bidirectional microphones; and using matrix systems, which electrically provide sum and difference signals from the left and right channels (these can be manipulated later for the desired effect).

The basic coincident technique was developed in the 1930s (along with the first stereo recordings) by English engineer Alan Blumlein.[14] Blumlein used two figure 8 pattern ribbon microphones mounted so that their pattern lobes were at right angles (90°) to each other, as shown in Fig. 16-104. The stereo effect is produced

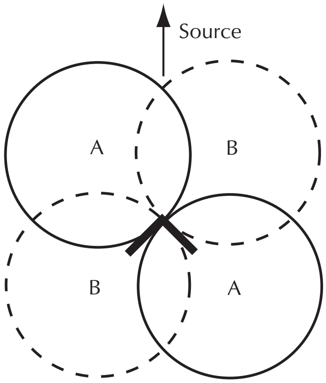


Figure 16-109. Coincident microphone technique using two bidirectional microphones.

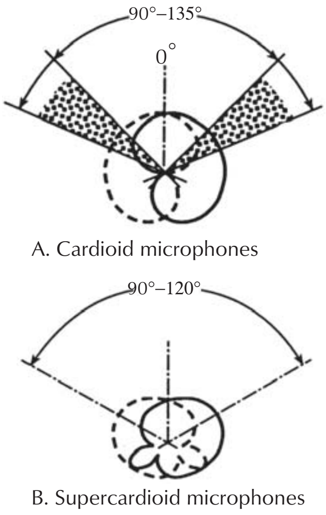
primarily by the difference in amplitude generated in the two microphones by the sound source. A sound on the right generates a larger signal in microphone B than in microphone A. A sound directly in front produces an equal signal in both microphones, and a sound on the left produces a larger signal in microphone A than in microphone B. The same process takes place with spaced omnidirectional microphones, but because of the spacing, there is also a time delay between two signals (comb filter effect). It can also produce a loss in gain and unpleasant sound if the two channels are combined into a single monosignal. Since the coincident microphone has both its transducers mounted on the same vertical axis, the arrival time is identical in both channels, reducing this problem to a large degree.

Modern coincident microphones often use cardioid or hypercardioid patterns. These patterns work as well as the figure 8 pattern microphones in producing a stereo image, but they pick up less of the ambient hall sound.

Probably the strongest virtue of the coincident microphone technique is its simplicity under actual working conditions. Just place the microphone in a central location that gives a good balance between the musicians and the acoustics of the hall. It is this simplicity that makes coincident microphones a favorite of broadcast engineers recording (or transmitting) live symphonic concerts.

16.7.2 XY Stereo Technique

The XY technique uses two identical directional microphones that, in relation to the recording axis, are arranged at equal and opposed offset angles. The leftward pointing X microphone supplies the L signal directly, and the rightward pointing Y microphone supplies the R signal, Fig. 16-110. The stereophonic properties depend on the directional characteristics of the microphones and the offset angle.



A. Cardioid microphones

B. Supercardioid microphones

Figure 16-110. XY stereo technique patterns.

One property specific to a microphone system is the recording angle, which defines the angle between the center axis (symmetry axis of the system) and the direction where the level differences between the L and R define the angular range of sound incidence where regular stereophonic reproduction is obtained. In most cases there is another opening for backward sound reception besides the recording angle for frontal sound pick-up.

Another important aspect concerns the relationship between the sound incidence angle and the stereophonic reproduction angle. As both XY and M/S recording techniques supply pure intensity cues, a relationship can be applied that relates the reproduction angle to the level difference of the L and R signals for the standard listening configuration based on an equilateral triangle, Fig. 16-111. This relationship is shown in Fig. 16-112 and is valid at frequencies between 330 and 7800 Hz within ±3°. The level difference is plotted on the horizontal axis and the reproduction angle can be read on the vertical scale. A 0° reproduction angle means localization at the center of the stereo base, and 30° means localization at one of the loudspeakers.

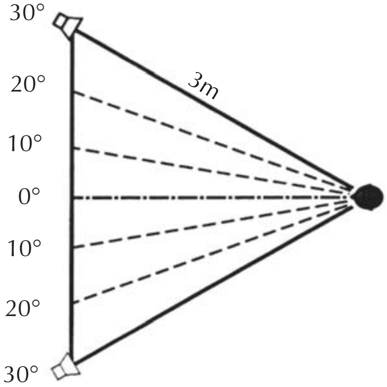


Figure 16-111. Standard listening configuration.

Fig. 16-113 shows the XY properties of wide-angle cardioids. The lower graph illustrates that the stereo image does not cover the full base width but is rather limited to some 20° at best. The recording angle can be altered between 90° and 120°. In-phase reproduction with correct side direction is maintained for all angles of sound incidence. The downward bend of the curves

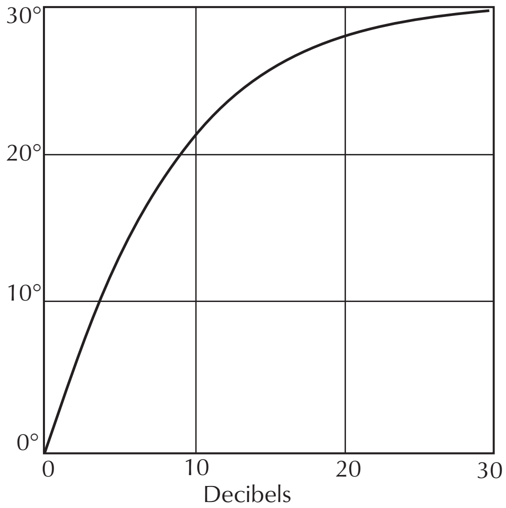


Figure 16-112. Stereophonic localization.

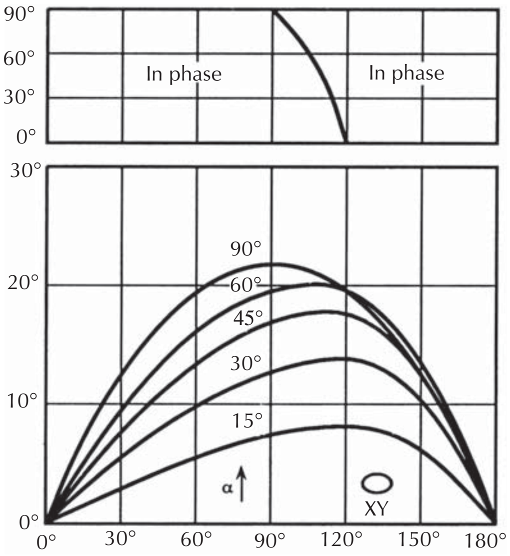


Figure 16-113. XY properties of wide-angle cardioids.

indicates that the stereo image is affected by geometric compression effects.

Fig. 16-114 shows the XY properties of cardioids. The recording angle can be altered between 90° and 180°. Again, in-phase reproduction at the correct side is maintained for all directions of sound incidence. As all reproduction curves touch the upper edge of the graph, full stereo width is available at all offset angles. The individual curvatures indicate deviations from the ideal geometrical reproduction performance which would be represented by a straight line. Downward bends indicate image compression at the base edges, whereas upward bends indicate expansion. The lower graph shows that compression effects occur at frontal sound pick-up for offset angles above 30°, whereas expansion occurs at smaller offset angles. Best reproduction linearity is performed at offset angles around 30°. The reproduction of back sound is always affected by angular compression. Extreme compression effects occur where the curves touch the upper edge of the graph. The reproduction is then clustered at one of the loudspeakers.

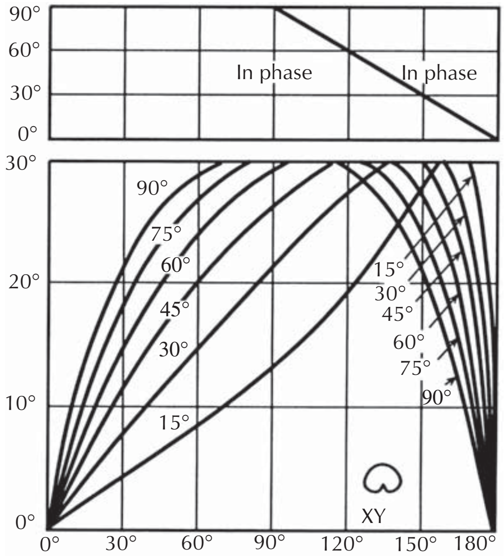


Figure 16-114. XY properties of cardioids.

16.7.3 The ORTF Technique

A variation on the basic XY coincident technique is the *ORTF technique*. The initials ORTF stand for *Office de Radiodiffusion Television Francais,* the French government radio network that developed this technique. The ORTF method uses two cardioid microphones spaced 7 inches (17 cm) apart and facing outward with an angle of 110° between them, Fig. 16-115. Because of the spacing between the transducers, the ORTF method does not have the time-coherence properties of M/S or XY micing.

16.7.4 The M/S Stereo Technique

The M/S technique employs a mid (M) cartridge that directly picks up the mono sum signal, and a side (S) cartridge that directly picks up the stereo difference signal (analogous to the broadcast stereo subcarrier

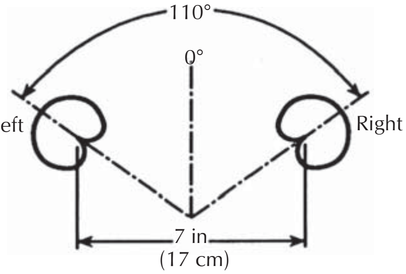


Figure 16-115. ORTF microphone technique.

Figure 16-116. Shure VP88 stereo condenser microphone. Courtesy Shure Incorporated.

Figure 16-117. AKG C422 stereo coincident microphone. Courtesy AKG Acoustics, Inc.

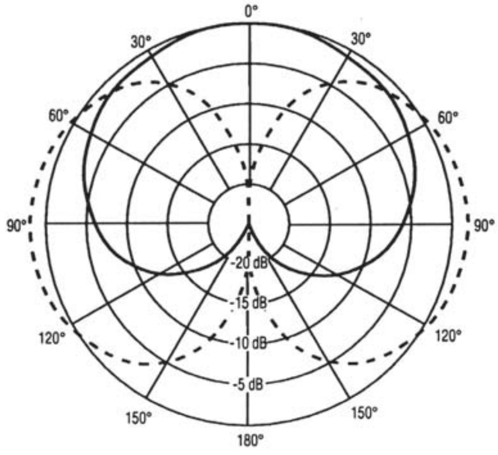


Figure 16-118. MS Microphone Pickup Patterns.

modulation signal). Although two individual microphones may be used, single-unit M/S microphones are more convenient and generally have closer cartridge placement. Fig. 16-116, a Shure VP88, and Fig. 16-117, an AKG C-422, are examples of M/S microphones.

Fig. 16-118 indicates the pickup patterns for a typical M/S microphone configuration. The mid cartridge is oriented with its front (the point of greatest sensitivity) aimed at the center of the incoming sound stage. A cardioid (unidirectional) pattern as shown is often chosen for the mid cartridge, although other patterns may also be used. For symmetrical stereo pickup, the side cartridge must have a side-to-side facing bidirectional pattern (by convention, the lobe with the same polarity as the front mid signal aims 90° to the left, and the opposite polarity lobe to the right).

In a stereo FM or television receiver, the mono sum baseband signal and the stereo difference subcarrier signal are demodulated and then decoded, using a sum-and-difference matrix, into left and right stereo audio signals. Similarly, the mid (mono) signal and the side (stereo difference) signal of the MS microphone may be decoded into useful left and right stereo signals.

The mid cartridge signal's relation to the mono sum signal, and the side cartridge signal's relation to the stereo difference signal, can be expressed simply by

$$M = \frac{1}{2(L+R)} \tag{16-28}$$

$$S = \frac{1}{2(L-R)} \tag{16-29}$$

Solving for the left and right signals,

$$L = M+S \tag{16-30}$$

$$R = M-S \tag{16-31}$$

Therefore, the left and right stereo signals result from the sum and difference, respectively, of the mid and side signals. These stereo signals can be obtained by processing the mid and side signals through a sum and difference matrix, implemented with transformers, Fig. 16-119, or active circuitry. This matrix may be external to the M/S microphone or built-in.

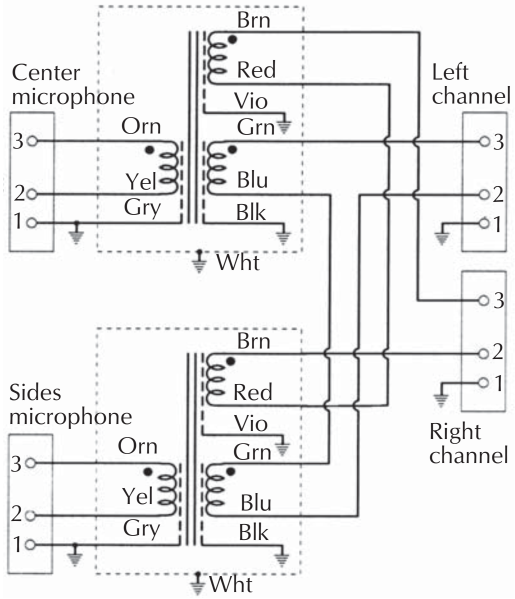


Figure 16-119. Transformer sum and difference matrix for M/S microphones.

In theory, any microphone pattern may be used for the mid signal pickup. Some studio M/S microphones provide a selectable mid pattern. In practice, however, the cardioid mid pattern is most often preferred in M/S microphone broadcast applications.

The AKG C422 shown in Fig. 16-120 is a studio condenser microphone that has been specially designed for sound studio and radio broadcasting. The microphone head holds two twin diaphragm condenser capsules elastically suspended to protect against handling noise.

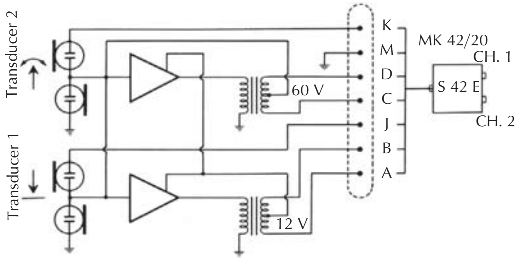


Figure 16-120. Schematic of the AKG C422 coincident microphone. Courtesy AKG Acoustics, Inc.

The wire-mesh grille is differently colored at the two opposing grille sides (light is the front grille side; dark is the rear grille side), thereby allowing the relative position of the two systems to be visually checked. The entire microphone can be rotated 45° about the axis to allow quick and exact changeover from 0° for M/S to 45° (for XY stereophony) even when the microphone is rigidly mounted. The upper microphone cartridge can be rotated 180° with respect to the lower one. A scale on the housing adjustment ring and an arrow-shaped mark on the upper system allows the included angle to be exactly adjusted. In sound studio work and radio broadcasts, it is often necessary to recognize the respective positions of the two systems from great distances; therefore, two light-emitting diodes with a particularly narrow light-emitting angle are employed. One is mounted in the upper (rotatable) housing, and the other in the lower (nonrotatable) housing. To align the heads, simply have the units rotated until the light-emitting diode is brightest on the preferred axis.

Enclosed within the microphone shaft are two separate field-effect transistor preamplifiers, one for each channel. The output level of both channels may be simultaneously lowered by 10 dB or 20 dB.

The C422 is connected to an S42E remote-control unit that allows any one of nine polar patterns to be selected for each channel. Because of noiseless selection, polar pattern changeover is possible even during recording.

Each channel of the microphone incorporates two cardioid diaphragms facing 180° of each other (back to back), Fig. 16-120. Note the 12 Vdc phantom power for the electronics and the 60 Vdc phantom power for the lower transducer (1), insuring that transducer 1 is always biased on. This transducer has a positive output for a positive pressure. The second or upper transducer is connected to pin K, which through the S42E, has nine switchable voltages between 0 Vdc and 120 Vdc. When the voltage at K is 60 Vdc, the output of transducer 2 is 0 (60 Vdc on either side of it), so the microphone output is cardioid.

When the voltage at K is 120 Vdc, transducer 2 is biased with 60 Vdc of an opposite polarity from transducer 1 so the output is 180° out of polarity, the mixed output being a figure 8 pattern.

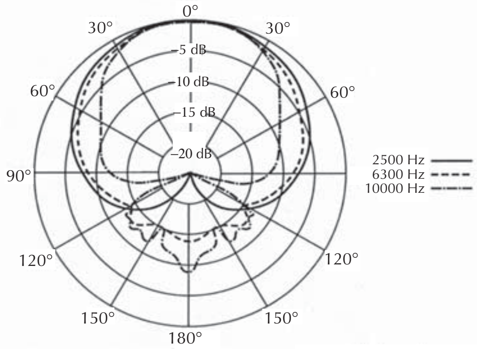
When the voltage at K is 0 Vdc, transducer 2 has a 60 Vdc bias on it with the same polarity as transducer 1. Because the transducers face in opposite directions, when these two outputs are combined, an omnidirectional pattern is produced.

By varying the voltage on K between 0 Vdc and 120 Vdc, various patterns between a figure 8 and an omnidirectional pattern can be produced.

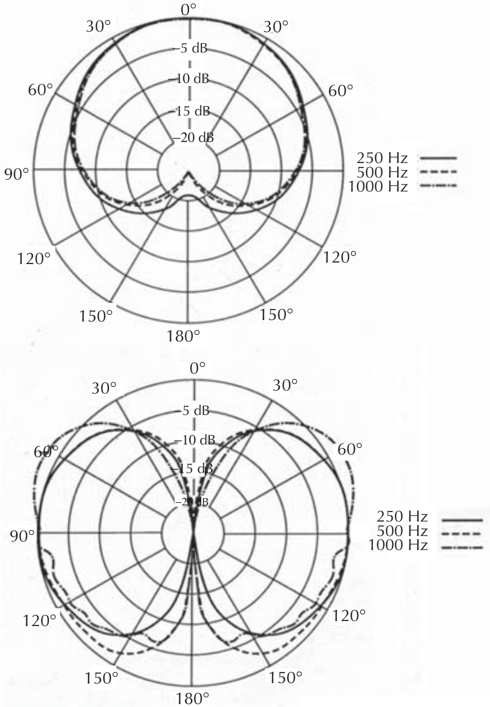
The Shure VP88 stereo microphone, Fig. 16-116, also employs a switchable pattern. Fig. 16-121 shows the polar response of the mid capsule and the side capsule.

The left and right stereo signals exhibit their own equivalent pickup patterns corresponding to, respectively, left-forward-facing and right-forward-facing microphones. Fig. 16-122 shows the relative levels of the mid and side microphones and the stereo pickup pattern of the Shure VP88 microphone in the L position with the bidirectional side pattern maximum sensitivity 6 dB lower than the maximum mid sensitivity. The small rear lobes of each pattern are 180° out of polarity with the main front lobes. For sound sources arriving at 0° the left and right output signals are equal, and a center image is reproduced between the loudspeakers. As the sound source is moved off-axis, an amplitude difference between left and right is created, and the loudspeaker image is moved smoothly off-center in the direction of the higher amplitude signal.

When the mid (mono) pattern is fixed as cardioid, the stereo pickup pattern can be varied by changing the side level relative to the mid level. Fig. 16-123 shows an M/S pattern in the M position with the side level 1.9 dB lower than the mid level. Fig. 16-124, position H, increases the side level to 1.6 dB higher than the mid level. The three resultant stereo patterns exhibit pickup angles of 180°, 127°, and 90°, respectively. The incoming sound angles, which will create left,



A. Mid-channel pattern



B. Side channel pattern

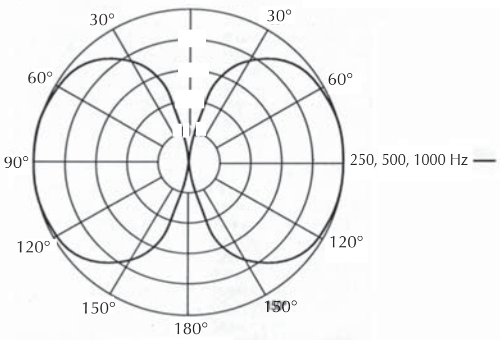
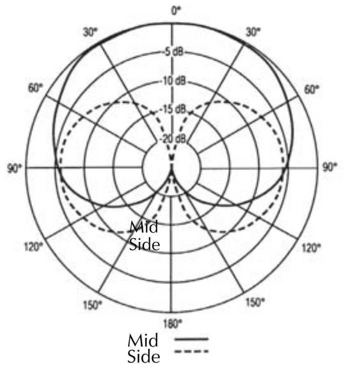
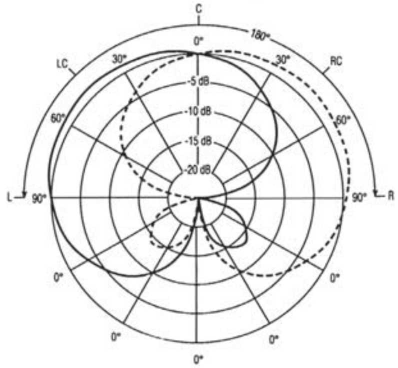


Figure 16-121. Polar response of the Shure VP88 M/S microphone. Courtesy Shure Incorporated.



A. Relative levels of the mid and side microphones.



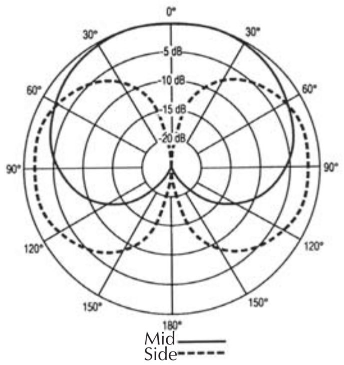
B. Pickup pattern of the system.

Figure 16-122. Stereo pickup pattern of the Shure VP88 in the L position. Courtesy Shure Incorporated.

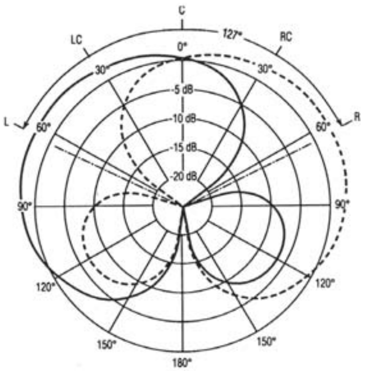
left-center, center, right-center, and right images, are also shown. Note the changes in the direction of the stereo patterns and the size of their rear lobes.

Taking the directional properties of real microphones into consideration, it becomes clear that the M/S technique provides a higher recording fidelity than the XY technique. There are at least three reasons for this:

1. The microphones in an XY system are operated mainly at off-axis conditions, especially at larger offset angles. The influence of directivity imperfections is more serious than with MS systems, where the M microphone is aimed at the performance center. This is illustrated by Fig. 16-125.
2. The maximum sound incidence angle for the microphone is only half that of the X and Y microphones, although the covered performance area is the same for all microphones. This area is symmetrically picked up by the M microphone, but unsymmetrically by the X and Y microphones. The M/S system can supply the more accurate monophonic (M) signal in comparison with the XY system.



A. Relative levels of the mid and side microphones.



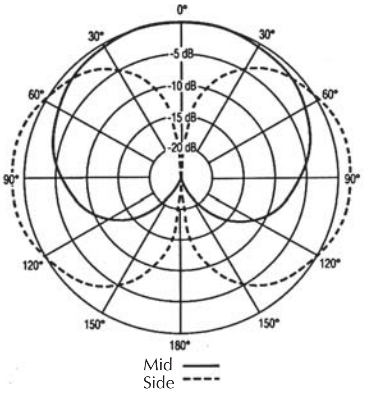
B. Pickup pattern of the system.

Figure 16-123. Stereo pickup pattern of the Shure VP88 in the M position. Courtesy Shure Incorporated.

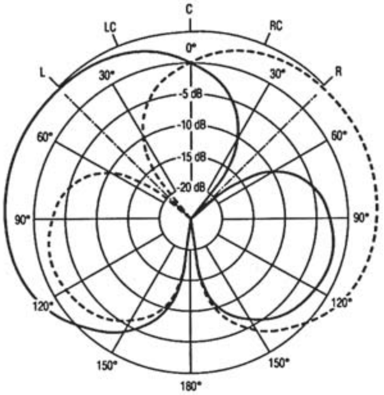
3. The M/S system picks up the S signal with a bidirectional microphone. The directivity performance of this type of microphone can be designed with a high degree of perfection, so errors in the S signal can be kept particularly small for all directions of sound incidence. The M/S system can supply a highly accurate side (S) signal.
4. In the M/S technique, the mono directivity does not depend on the amount of S signal applied to create the stereophonic effect. If recordings are made in the M/S format, a predictable mono signal is always captured. On the other hand, the stereophonic image can be simply influenced by modifying the S level without changing the mono signal. This can even be done during postproduction.

16.7.5 The Stereo Boom Microphone

Microphones for stereophonic television broadcasting have their own special problems. The optimal distance



A. Relative levels of the mid and side microphones.



B. Pickup pattern of the system.

Figure 16-124. Stereo pickup pattern of the Shure VP88 in the H position. Courtesy Shure Incorporated.

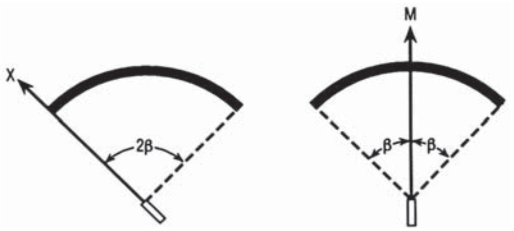


Figure 16-125. Exposure of X and M microphones to 0° pickup.

from the TV screen is considered to be five times the screen diagonal, at which distance the line structure of the TV image can no longer be resolved by the human eye. The resulting minimum observer distance is therefore 11 ft (3.3 m) and the two loudspeakers should be at least 12.5 ft (3.8 m) apart. This is certainly not realistic for television viewing.

The sound engineer must take into account that the reproduction will be through loudspeakers right next to the TV screen as well as through hi-fi equipment. If, for instance, the full base width is used during the sound recording of a TV scene, an actor appearing at the right edge of the picture will be heard through the right loudspeaker of the hi-fi system and sound as if he is far to the right of the TV set. This will result in an unacceptable perception of location.

The viewer must be able to hear realistically the talker on the TV screen in the very place where the viewer sees the talker. To achieve this goal, German television proposed to combine a unidirectional microphone for the recording of the actors with a figure 8microphone for the recording of the full stereophonic basic width.

This recording technique utilizes a figure-eight microphone suspended from the boom in such a way that it maintains its direction when the boom is rotated while a second microphone with unidirectional pattern is mounted on top and follows the movement of an actor or reporter, Fig. 16-126. To make sure that the directional characteristic of the moving microphone does not have too strong an influence and that a slight angular error will not lead to immediately perceptible directional changes, the lobe should be somewhat wider as with the customary shotgun microphones in use today.

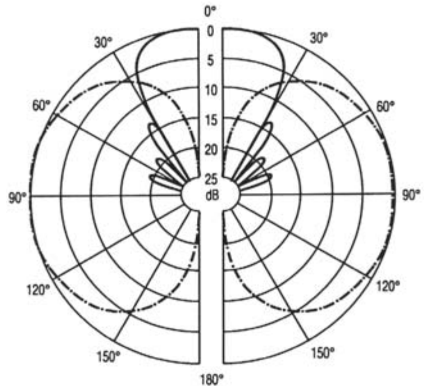


Figure 16-126. Stereo boom microphone using Sennheiser MKH 30 and MKH 70 microphones. Courtesy Sennheiser Electronic Corporation.

It is now possible to produce a finished stereo soundtrack on location by positioning the microphones in the manner of an M/S combination. The level of the figure 8 microphone can be lowered and used only for the recording of voices outside of the picture, ambience and music. This microphone should always remain in a fixed position and its direction should not be changed. The S-signal generated in this way must be attenuated to such a degree that the M-signal microphone will always remain dominant. This microphone is the one through which the actors pictured on the screen will be heard.

16.7.6 SASS Microphones

The Crown® SASS-P MK II or Stereo Ambient Sampling System™, Fig. 16-127, is a patented, mono-compatible, near-coincident array stereo condenser microphone using PZM technology.

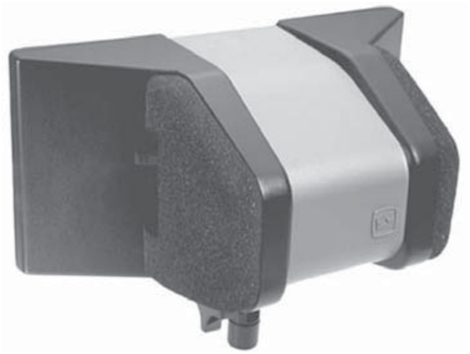


Figure 16-127. Crown® SASS-P MK II stereo microphone. Courtesy of Crown International, Inc.

The SASS uses two PZMs mounted on boundaries (with a foam barrier between them) to make each microphone directional. Another Crown model, SASS-B, is a similarly shaped stereo boundary mount for Brüel & Kjaer 4006 microphones and is used for applications requiring extremely low noise.

Controlled polar patterns and human-head-sized spacing between capsules create a focused, natural stereo image with no hole-in-the-middle for loudspeaker reproduction, summing comfortably to mono if required.

The broad acceptance angle (125°) of each capsule picks up ambient sidewall and ceiling reflections from the room, providing natural reproduction of acoustics in good halls and ambient environments. This pattern is consistent for almost ±90° vertical.

A foam barrier/baffle between the capsules shapes the pickup angle of each capsule toward the front, limiting overlap of the two sides at higher frequencies. Although the microphone capsules are spaced a few centimeters apart, there is little phase cancellation when both channels are combined to mono because of the shadowing effect of the baffle. While there are phase differences between channels, the extreme amplitude differences between the channels caused by the baffle, reduce phase cancellations in mono up to 20 kHz.

The SASS has relatively small boundaries. However, it has a flat response down to low frequencies because there is no 6 dB shelf as in standard PZM microphones (see Section 16.6.1.2.3). The flat response is attained because the capsules are omnidirectional below 500 Hz, and their outputs at low frequencies are equal in level, which, when summed in stereo listening, causes a 3 dB rise in perceived level. This effectively counteracts one half of the low frequency shelf normally experienced with small boundaries.

In addition, when the microphone is used in a reverberant sound field, the effective low-frequency level is boosted another 3 dB because the pattern is omnidirectional at low frequencies and unidirectional at high frequencies. All of the low-frequency shelf is compensated, so the effective frequency response is uniform from 20 Hz–20 kHz. Fig. 16-128 is the polar response of the left channel (the right channel is the reverse of the left channel).

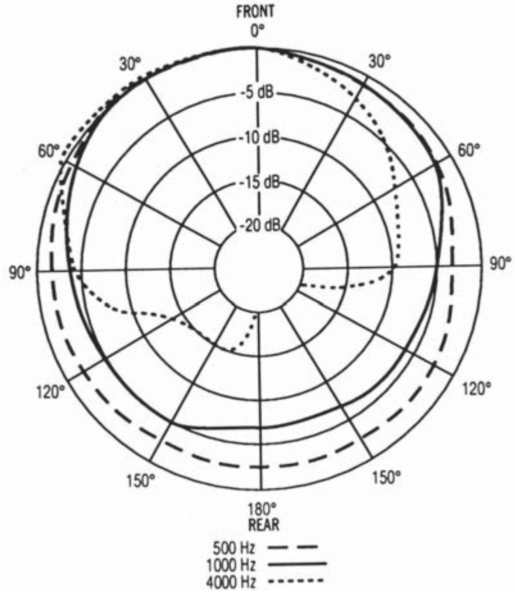


Figure 16-128. SASS-P MK II polar response, of the left channel. 0° sound incidence is perpendicular to the boundary. The right channel is a mirror image of the left channel. Courtesy Crown International.

16.7.7 Surround Sound Microphone System

16.7.7.1 Schoeps 5.1 Surround System

The Schoeps 5.1 surround system consists of the KFM 360 sphere microphone, two figure 8 microphones with suspension and the DSP-4 KFM 360 processor, Fig. 16-129.

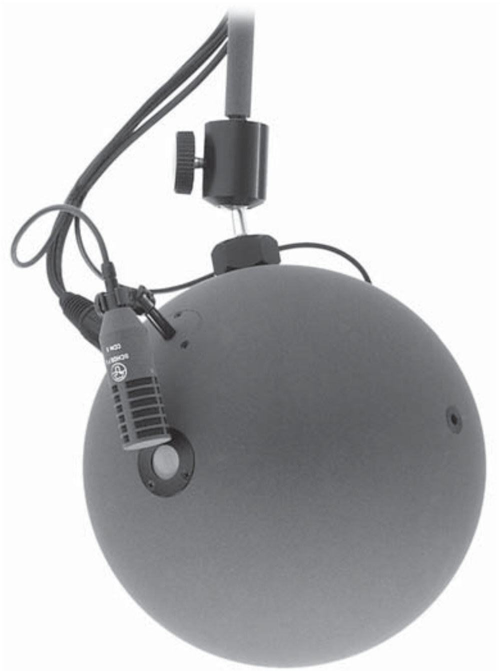


Figure 16-129. Schoeps KFM 360 sphere microphone. Courtesy Schoeps GmbH.

The central unit in this system is the sphere microphone KFM 360. It uses two pressure transducers and can, even without the other elements of the system, be used for stereophonic recording. Its recording angle is about 120°, permitting closer micing than a standard stereo microphone. The necessary high-frequency boost is built into the processor unit.

Surround capability is achieved through the use of two figure 8 microphones, which can be attached beneath the pressure transducers by an adjustable, detachable clamp system with bayonet-style connectors. The two microphones should be aimed forward.

The DSP-4 KFM 360 processor derives the four corner channels from the microphone signals. A center channel signal is obtained from the two front signals, using a special type of matrix. An additional channel carries only the low frequencies, up to 70 Hz. To avoid perceiving the presence of the rear loudspeakers, it is possible to lower the level of their channels, to delay them and/or to set an upper limit on their frequency response, Fig. 16-130.

The front stereo image width is adjustable and the directional patterns of the front-facing and rear-facing pairs of virtual microphones can be chosen independently of one another.

The processor unit offers both analog and digital inputs for the microphone signals. In addition to providing gain, it offers a high-frequency emphasis for the built-in pressure transducers as well as a low-frequency boost for the figure 8s.

As with M/S recording, matrixing can be performed during post-production in the digital domain.

The system operates as follows: the front and rear channels result from the sum (front) and difference (rear) of the omnidirectional and figure 8 microphones on each side, respectively, Fig. 16-131. The four resulting virtual microphones that this process creates will seem to be aimed forward and backward, as the figure 8s are. At higher frequencies they will seem to be aimed more toward the sides (i.e., apart). Their directional pattern can be varied, anywhere from omnidirectional to cardioid to figure 8. The pattern of the two rear-facing virtual microphones can be different from that of the two forward-facing ones. Altering the directional patterns alters the sound as well, in ways that are not possible with ordinary equalizers. This permits a flexible means of adapting to a recording situation—to the acoustic conditions in the recording space—and this can even be done during postproduction, if the unprocessed microphones signals are recorded.

This four-channel approach yields a form or surround reproduction without a center channel—something that is not what everybody requires.

16.7.7.2 Holophone® H2-PRO Surround Sound System

The elliptical shape of the Holophone® H2-PRO emulates the characteristics of a human head, Fig. 16-132. Sound waves bend around the H2-PRO as they do around the head providing an accurate spatiality, audio imaging, and natural directionality. Capturing the directionality of these soundwaves translates into a very realistic surround sound experience. The total surface area of the eight individual elements combines with the spherical embodiment of the H2-PRO to capture the acoustic textures required for surround reproduction, Fig. 16-133. The embodiment acts as an acoustic lens capturing lows and clean highs.

A complete soundfield can be accurately replicated without the use of additional microphones—a simple point-and-shoot operation. The Holophone H2-PRO is

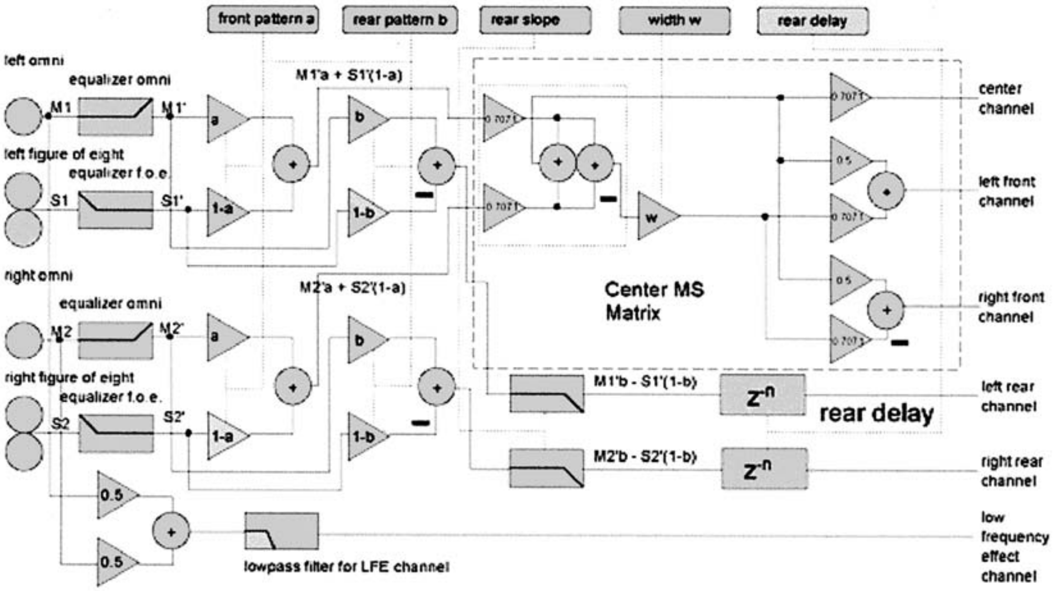

front pattern a
rear pattern b
rear slope
width w
rear delay
left omni
equalizer omni
M1
M1'
left figure of eight
equalizer f.o.e.
S1
S1'
a
1-a
M1'a + S1'(1-a)
b
1-b
0.707
w
Center MS Matrix
0.5
center channel
left front channel
right front channel
right omni
equalizer omni
M2
M2'
right figure of eight
equalizer f.o.e.
S2
S2'
M2'a + S2'(1-a)
M1'b - S1'(1-b)
Z^-n
rear delay
left rear channel
M2'b - S2'(1-b)
right rear channel
0.5
lowpass filter for LFE channel
low frequency effect channel

Figure 16-130. Schoeps DSP-4 KFM 360 processor. Courtesy Schoeps GmbH.

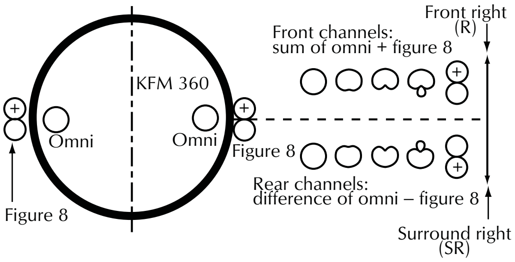

KFM 360
Omni
Omni
Figure 8
Figure 8
Front right (R)
Front channels: sum of omni + figure 8
Rear channels: difference of omni – figure 8
Surround right (SR)

Figure 16-131. Derivation of the right (R) and right surround signals (SR) of the Schoeps 5.1 Surround System. Courtesy Schoeps GmbH.

capable of recording up to 7.1 channels of discrete surround sound. It terminates in eight XLR microphone cable ends (Left, Right, Center, Low Frequency, Left Surround, Right Surround, Top, and Center Rear). These co-relate to the standard 5.1 channels and add a top channel for formats such as IMAX and a center rear channel for extended surround formats such as Dolby EX, DTS, ES, and Circle Surround. Because each microphone has its own output, the engineer may choose to use as many or as few channels as the surround project requires as channel assignments are discrete all the way from the recording and mixing process to final delivery. It is well suited for television broadcasters (standard TV, DTV, and HDTV), radio broadcasters, music producers and engineers, film location recording crews, and independent project studios.

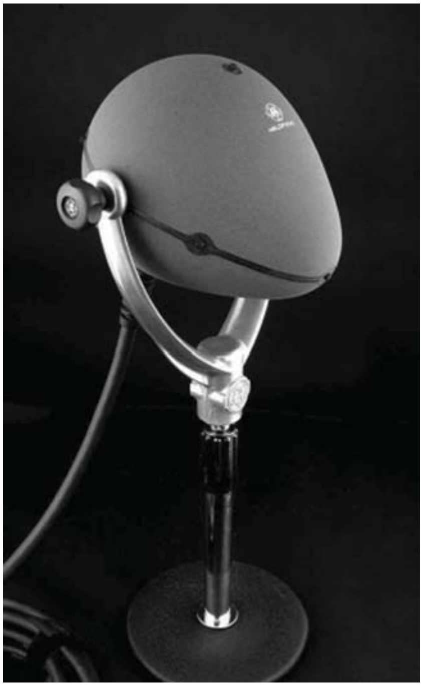


Figure 16-132. Holophone H2-PRO surround sound system. Courtesy Holophone®.

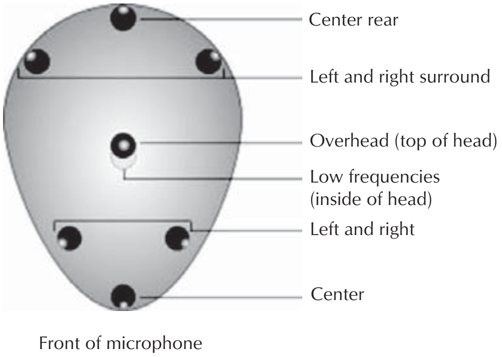


Figure 16-133. Location of the microphones on the H2-PRO head. Courtesy Holophone®.

16.7.7.3 Holophone H4 SuperMINI Surround Sound System

The H4 SuperMINI head, Fig. 16-134 contains six microphone elements, that translate to the standard surround sound loudspeaker configuration; L, R, C, LFE, LS, RS. The LFE collects low-frequency signals for the subwoofer. The six discrete channels are fed into a Dolby® Pro-Logic II encoder which outputs the audio as a stereo signal from a stereo mini-plug to dual XLRs, dual RCAs, or dual mini-plugs. The left and right stereo signals can then be connected to the stereo audio inputs of a video camera or stereo recorder. The encoded signal is recorded onto the media in the camera or recorder and the captured audio can be played back in full 5.1-channel surround over any Dolby® Pro Logic II–equipped home theatre system.The material can be edited and the audio can be decoded via a Dolby® Pro Logic II Decoder and then brought into an NLE including Final Cut or iMovie, etc. The stereo recording can also be broadcast directly through the standard infrastructure. Once it is received by a home theatre system, containing a Dolby® Pro-Logic II or any compatible decoder, the six channels are completely unfolded to their original state. Where no home theatre receiver is detected, the signal will simply be heard in stereo. The SuperMINI has additional capabilities that include an input for an external, center-channel-placed shotgun or lavalier microphone to enhance sonic opportunity options and features an audio zoom button that increases the forward bias of the pickup pattern. It also includes virtual surround monitoring via headphones for real-time on-camera 3D audio monitoring of the surround field.

HOLOPHONE
H4 SUPERMINI

Figure 16-134. Holophone H4 SUPERmini Surround SSound System. Courtesy Holophone®.

16.8 Microphones for Binaural Recording

16.8.1 Artificial Head Systems

Human hearing is capable of selecting single sounds from a mixture of sounds while suppressing the unwanted components (the cocktail party effect). This is done in the listener's brain by exploiting the ear signals as two spatially separated sound receivers in a process frequently referred to as *binaural signal processing*. A simple test will verify this statement: when listening to a recording of several simultaneous sound events recorded by a single microphone, the individual sources cannot be differentiated.

Two spaced microphones or more elegant multielement spatially sensitive microphones, such as a stereo coincident microphone, have been used to capture the spatial characteristics of sounds, but they have frequently been deficient when compared to what a person perceives in the same environment. This lack of realism is attributed to absence of the spectrum modification inherent in sound propagation around a person's head and torso and in the external ear—i.e., the transfer function of the human and the fact that the signals are kept separate until very late in the human analysis chain.

The acoustic transfer function of the human external ear is uniquely related to human body geometry. It is composed of four parts that can be modeled mathemati-

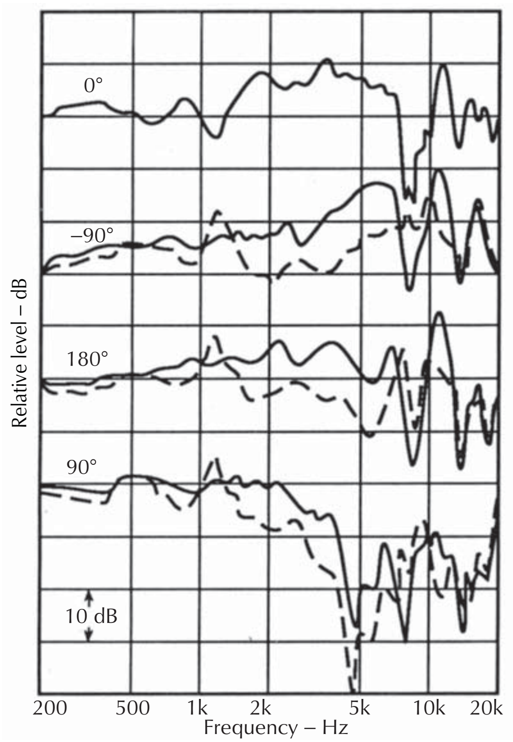


Figure 16-135. Transfer function of the left ear, measured 4 mm inside the entrance of the ear canal, for four angles of incidence (straight ahead, to the left, straight behind, and to the right).

cally, as shown in Fig. 16-135, or recreated by an artificial head system.[15,16,17]

Reflections and diffraction of sound at the upper body, the shoulder, the head, and the outer ear (pinna), as well as resonances caused by the concha and ear canal, are mainly responsible for the transfer characteristic. The cavum concha is the antechamber to the ear. The spectral shape of the external ear transfer function varies from person to person due to the uniqueness of people and the dimensions of these anatomical features. Therefore, both artificial heads and their mathematical models are based on statistical composites of responses and dimensions of a number of persons.

All of these contributions to the external ear transfer function are direction sensitive. This means that sound from each direction has its own individual frequency response. In addition, the separation of the ears with the head in between affects the relative arrival time of sounds at the ears. As a result the complete outer-ear transfer function is very complicated, Fig. 16-136, and can only be partially applied as a correction to the response of a single or even a pair of microphones. In the figure, the base of each arrow indicates reference SPL. The solid curves represent the free-field (direct sound) external ear transfer function, while the dashed curves represent the difference, at each direction, relative to frontal free-field sound incidence.

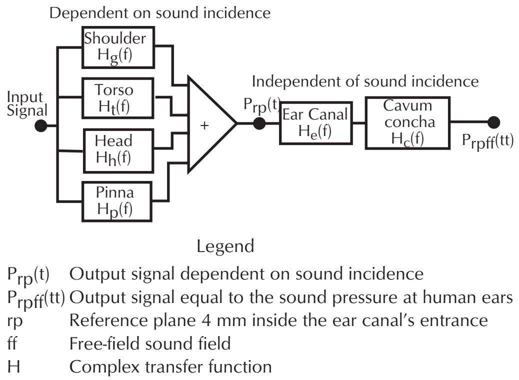

Dependent on sound incidence
Shoulder Hg(f)
Torso Ht(f)
Head Hh(f)
Pinna Hp(f)
Input Signal
+
Independent of sound incidence
Prp(t)
Ear Canal He(f)
Cavum concha Hc(f)
Prpff(tt)
Legend
Prp(t) Output signal dependent on sound incidence
Prpff(tt) Output signal equal to the sound pressure at human ears
rp Reference plane 4 mm inside the ear canal's entrance
ff Free-field sound field
H Complex transfer function

Figure 16-136. The human external-ear transfer function.

Artificial heads have been used for recording for some time. However, the latest heads and associated signal processing electronics have brought the state of the art close to in the ear (ITE) recording, which places microphones in human ears.

The KU 100 measurement and recording system by Georg Neumann GmbH in Germany is an example of a high-quality artificial head system, Fig. 16-137. Originally developed by Dr. Klaus Genuit and his associates at the Technical University of Aachen, the artificial head, together with carefully designed signal processing equipment, provides binaural recording systems that allow very accurate production of spatial imaging of complex sound fields.

The head is a realistic replica of a human head and depends on a philosophy of sound recording and reproduction—namely, that the sound to be recreated for a listener should not undergo two transfer functions, one in the ears of the artificial head and one in the ears of the listener.

Fig. 16-138 is a block diagram of a head microphone and recording system. A high-quality microphone is mounted at the ear canal entrance position on each side of the head. Signals from each microphone pass through diffuse-field equalizers in the processor and are then available for further use in recording or reproduction. The diffuse-field equalizer is specifically tuned for the head to be the inverse of the frontal diffuse-field transfer function of the head. This signal is then recorded and can be used for loudspeaker playback and for measurement. The headphone diffuse-field equalizers in the Reproduce Unit yield a linear diffuse-field

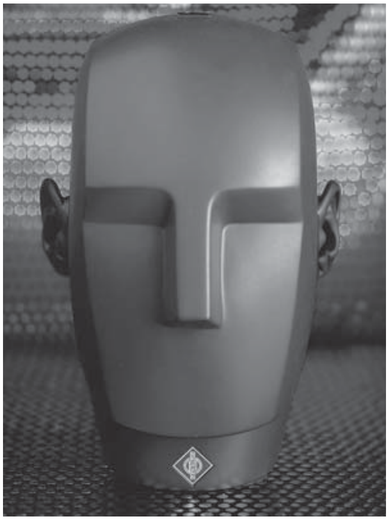


Figure 16-137. Georg Neumann KU 100 dummy head. Courtesy Georg Neumann GmbH.

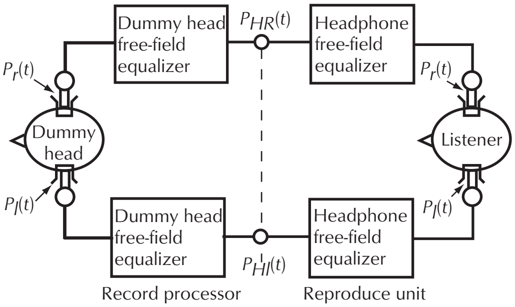

Dummy head free-field equalizer
$P_{HR}(t)$
Headphone free-field equalizer
$P_r(t)$
$P_r(t)$
Dummy head
Listener
$P_l(t)$
$P_l(t)$
Dummy head free-field equalizer
$P_{Hl}(t)$
Headphone free-field equalizer
Record processor
Reproduce unit

Figure 16-138. Block diagram of a dummy head binaural stereo microphone system.

transfer function of the headphone, so the sound pressures presented at the entrance of the listener's ear canals will duplicate those at the entrance of the head's ear canals.

Diffuse-field equalization is suitable for situations where the source is at a distance from the head. For recordings close to a sound source or in a confined space, such as the passenger compartment of an automobile, another equalization called *independent of direction* (ID) is preferred. The equalization is internal in the head in Fig. 16-138.

Signals $P_{Hr}(t)$ and $P_{Hl}(t)$ from the heads can be recorded and used directly for loudspeaker playback, analysis, or playback through headphones. As a recording tool this method can surpasse many other recording techniques intended for loudspeaker reproduction. The full benefits of spatial imaging can be heard and enjoyed with earphone playback as well as with high quality loudspeakers.

The heads are constructed of rugged fiberglass. The microphones can be calibrated by removal of the detachable outer ears and applying a pistonphone. Preamplifiers on the microphones provide polarization and have balanced transformerless line drivers. A record processor and modular unit construction provides dc power to the dummy head and act as the interface between the head and the recording medium or analysis equipment. The combination of low noise electronics and good overload range permits full use of the 135 dB dynamic range of the head microphones and 145 dB with the 10 dB attenuator switched in.

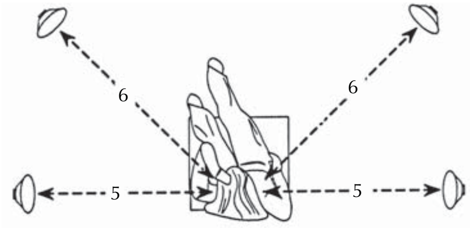
For headphone playback, a reproduce unit provides an equalized signal for the headphones that produces earcanal entrance sound signals that correspond to those at the corresponding location on the artificial head.

An important parameter to consider in any head microphone recording system is the dynamic range available at this head signal output. For example, the canal resonance can produce a sound pressure that may exceed the maximum allowed on some ear canal-mounted microphones.

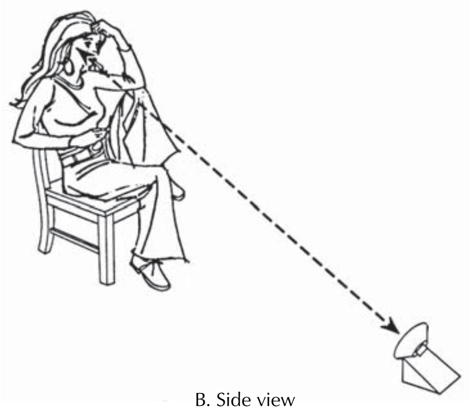
16.8.2 In the Ear Recording Microphones[18]

In-the-Ear (ITE™) recording and Pinna Acoustic Response (PAR™) playback represent a new-old approach to the recording of two channels of spatial images with full fidelity and their playback over two channels, Fig. 16-138. It is important that the loudspeakers are in signal synchronization and that they be placed at an angle so that the listener position is free of early reflections.

Low noise, wide frequency, and dynamic range probe microphones employing soft silicone probes are placed in the pressure zone of the eardrum of live listeners. This microphone system allows recording with or without equalization to compensate for the ear canal resonances while leaving the high-frequency comb filter spatial clues unaltered. The playback system consists of synchronized loudspeaker systems spaced approximately equal distances from the listener in the pattern shown in Fig. 16-139. Both left loudspeakers are in parallel, and both right loudspeakers are in paralle. However, the front and back loudspeakers are on individual volume controls. This is to allow balancing side-to-side and to adjust the front-to-back relative levels for each individual listener. The two front loud-



A. Top view



B. Side view

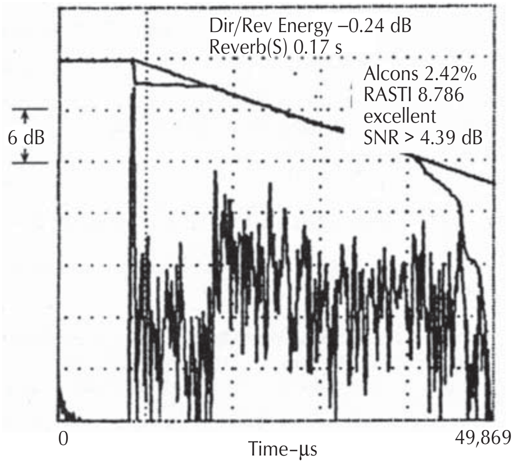
Figure 16-139. The loudspeaker arrangement used for PAR playback of ITE recordings. Courtesy Syn-Aud-Con.

speakers are used to provide hearing signals forward of the listener.

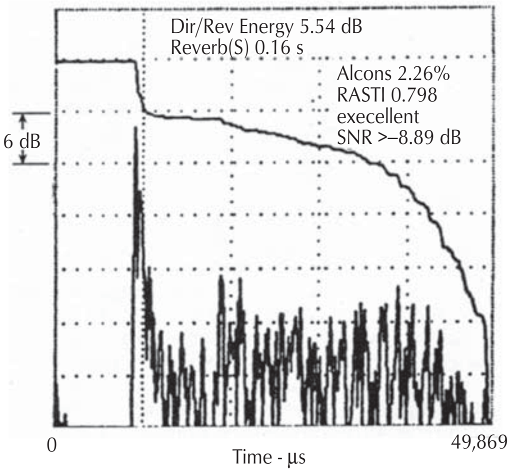
Fig. 16-140A shows an ETC made in a listening room ($L_D - L_R = 0.24$). Fig. 16-140B is the identical position measured with the ITE technique ($L_D - L_R = 5.54$). Note particularly the difference in $L_D - L_R$ for the two techniques. ITE recording and PAR playback allow a given listener to hear a given speech intelligibility environment as perceived by another person's head and outer ear configuration right up to the eardrum.

Recordings made using ITE microphones in two different people's ears of the same performance in the same seating area sound different. Playback over loudspeakers where the system is properly balanced for perfect geometry for one person may require as much as 10 dB different front to back balance for another person to hear perfect geometry during playback.

Since ITE recordings are totally compatible with normal stereophonic reproduction systems and can provide superior fidelity in a many cases, the practical use of ITE microphony would appear to be unlimited.



A. An ETC made in a listening room with a GenRAd 1/2 inch microphone ($L_o - L_R = 0.24$).



B. The identical position measured with the ITE technique ($L_D - L_R = 0.24$).

Figure 16-140. An ETC comparison of a measurement microphone and the ITE technique at the same position in the room. Courtesy Syn-Aud-Con.

16.9 USB Microphones

The computer has become a important part of sound systems. Many consoles are digital and microphones are connected directly to them. Microphones are also connected to computers through the USB input.

The audio-technica AT2020 USB cardioid condenser microphone, Fig. 16-141 is designed for computer-based recording. It includes a USB (Universal Serial Bus) digital output that is Windows and Mac compatible. The sample rate is 44.1 kHz with a bit depth of

16 bits. The microphone is powered directly from the 5 Vdc USB output.

The MXL.006 USB is a cardioid condenser microphone with a USB output that connects directly to a computer without the need for external mic preamps through USB 1.1 and 2.0, Fig. 16-142.

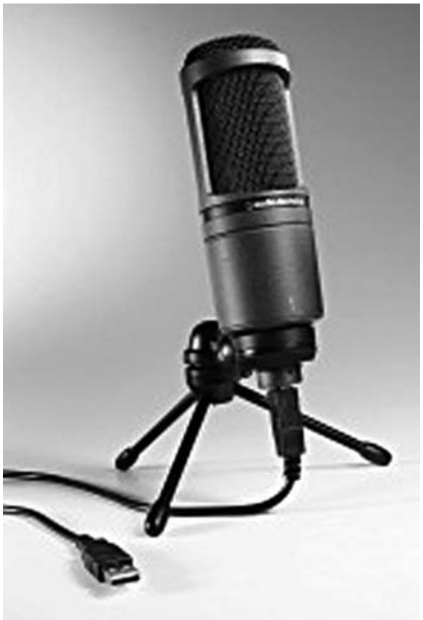


Figure 16-141. audio-technica AT2020 USB microphone. Courtesy Audio-Technica U.S., Inc.

Figure 16-142. MXL.006 USB microphone. Courtesy Marshall Electronics, Inc.

The analog section of the MXL.006 microphone features a 20 Hz–20 kHz frequency response, a large gold diaphragm, pressure-gradient condenser capsule, and a three-position, switchable attenuation pad with settings for Hi (0 dB), Medium (−5 dB), and Lo (−10 dB). The digital section features a 16-bit Delta Sigma A/D converter with a sampling rate of 44.1 kHz and 48 kHz. Protecting the instrument's capsule is a heavy-duty wire mesh grill with an integrated pop filter.

The MXL.006 includes a red LED behind the protective grill to inform the user that the microphone is active and correctly oriented. The MXL.006 ships with a travel case, a desktop microphone stand, a 10 ft USB cable, windscreen, an applications guide, and free downloadable recording software for PCs and Mac.

16.10 Wireless Communication Systems

Wireless communication systems are wireless microphones (*radio microphones*), Fig. 16-143, and a related concept, wireless intercoms. Often the same end user buys both the microphones and intercoms for use in television and radio broadcast production, film production, and related entertainment-oriented applications.

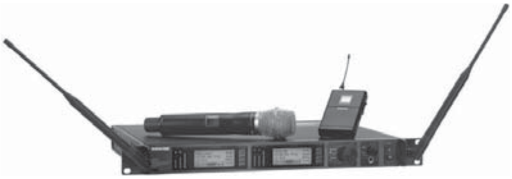


Figure 16-143. Shure UHF-R Wireless Microphone System. Courtesy Shure Incorporated.

Wireless microphone systems can be used with any of the preceding microphones discussed. Some wireless microphone systems include a fixed microphone cartridge while others allow the use of cartridges by various manufacturers.

A block diagram of a wireless microphone system is shown in Fig. 16-144. The sending end of a wireless microphone system has a dynamic, condenser, electret, or pressure zone microphone connected to a preamplifier, compressor, and a small transmitter/modulator and antenna.

The receiving end of the system is an antenna, receiver/discriminator, expander, and preamplifier, which is connected to external audio equipment.

In a standard intercom system, each person has a headset and belt pack (or equivalent), all interconnected by wires. Wireless intercoms are essentially identical in operation, only they use no cable between operators. Instead, each belt pack includes a radio transmitter and receiver. The wireless intercom user typically

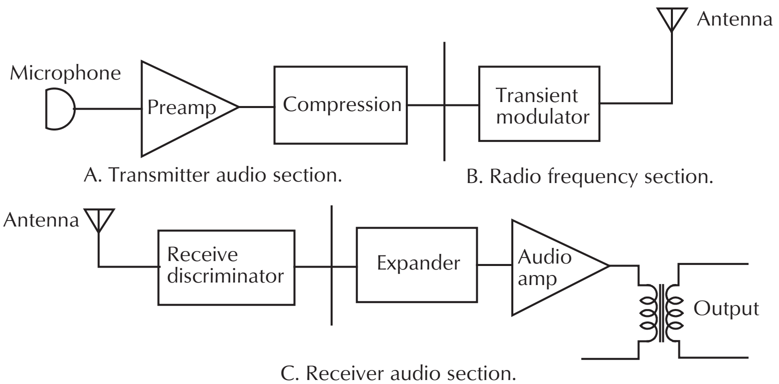


Figure 16-144. Wireless microphone transmitter section with built-in preamplifier, compressor, and transmitter, and the receiver with a built-in discriminator expander.

wears a headset (a boom microphone with one or two earpieces) and can simultaneously transmit on one frequency and receive on another. The wireless intercom transmitter is virtually identical to a wireless microphone transmitter, but the receiver is miniaturized so that it, too, can be conveniently carried around and operated with minimum battery drain.

Wireless microphones are widely used in television production. Handheld models (integral microphone capsule and transmitter) are used by performers "on camera," where they not only free the performer to walk around and gesture spontaneously, they also avoid the need for stage personnel to feed wires around cameras, props, etc. Lavalier models (small pocket-sized transmitters that work with lavalier or miniature "hidden" microphones) are used in game shows, soap operas, dance routines, etc., where they eliminate the need for boom microphones and further avoid visual clutter.

For location film production, as well as electronic news gathering (ENG) and electronic field production (EFP), wireless microphones make it possible to obtain usable "first take" sound tracks in situations where previously, postproduction dialogue looping was necessary.

In theatrical productions, wireless microphones free actors to speak and/or sing at normal levels through a properly designed sound-reinforcement system.

In concerts, handheld wireless microphones permit vocalists to move around without restriction, and without shock hazard even in the rain. Some lavalier models have high-impedance line inputs that accept electric guitar cords to create wireless guitars.

In all of the above applications where wireless microphones are used, in the studio or on location, a wireless intercom also is an invaluable communication aid between directors, stage managers, camera, lighting and sound crews, and security personnel. For cueing of talent and crews (or monitoring intercom conversations), economical receive-only units can be used. In sports production, wireless intercoms are used by coaches, spotters, players, production crews, and reporters. A major advantage is zero setup time. In critical stunt coordination, a wireless intercom can make the difference between a safe event or none at all. For more information on intercoms, refer to Chapter 43.

16.10.1 Criteria for Selecting a Wireless Microphone

There are a number of criteria that must be considered in obtaining a wireless microphone system suitable for professional use.[19,20] Ideally, such a system must work perfectly and reliably in a variety of tough environments with good intelligibility and must be usable near strong RF fields, lighting dimmers, and sources of electromagnetic interference. This relates directly to the type of modulation (standard frequency modulation or narrow-band frequency modulation), the operating frequency, high frequency (HF), very high frequency (VHF), ultrahigh frequency (UHF), the receiver selectivity, and so forth. The system must be very reliable and capable of operating at least five hours on one set of disposable batteries (or on one recharge if Ni-Cads are used).

16.10.1.1 Frequency Bands of Operation

Based on the FCC's reallocation of frequencies and the uncertainty of current and future allocations, some wireless manufacturers are offering systems that avoid the VHF and UHF bands completely. The ISM (industrial, science, and medicine) bands provide a unique alternative to the TV bands. By international agreement, all devices are low powered so there will never be any grossly high-powered RF interference potential. The 2.4 GHz band provides a viable alternative to traditional UHF bands, and as long as line of sight between transmitters and receivers is monitored users can easily get a 100 meter range. Another benefit of 2.4 GHz is that it can simplify wireless inventory for traveling shows. The same wireless frequencies are accepted worldwide, so there is no need to adhere to the country-specific frequency rules that severely complicate the situation for international tours. The same applies within the United States—the same frequencies work in all areas.

Currently wireless microphones are licensed on several frequencies, the most common being:

VHF low band (AM and FM)	25 to 50 MHz 72 to 76 MHz
FM broadcast (FM)	88 to 108 MHz
VHF high band (FM)	150 to 216 MHz
UHF (FM)	470 to 746 MHz 902 to 952 MHz

The VHF bands are seldom used anymore and can only be found on old equipment. The low band is in the noisiest radio spectrum and, because the wavelength is about 20 ft (6 m) it requires a long antenna (5 ft or 1.5 m). The VHF low band is susceptible to skip, which can be defined as external signals from a long distance away bouncing off the ionosphere back to earth, creating interference.

The VHF high band is more favorable than the low band. The ¼-wavelength antenna is only about 17 in (43 cm) long and requires little space. The VHF band has some penetration through buildings that can be advantageous and disadvantageous. It is advantageous in being able to communicate between rooms and around surfaces. It is disadvantageous in that transmission is not controlled (security), and outside noise sources can reach the receiver.

Most often the frequencies between 174 MHz and 216 MHz are used in the VHF band, corresponding to television channels 7 to 13. The VHF high band is free of citizens band and business radio interference, and any commercial broadcast stations that might cause interference are scheduled so you know where they are and can avoid them. Inherent immunity to noise is built in because it uses FM modulation. Better VHF high-band receivers have adequate selectivity to reject nearby commercial television or FM broadcast signals. If operating the microphone or intercom on an unused television channel—for instance Channel 7—protection might be required against a local television station on Channel 8. Another problem could be caused by an FM radio station. If a multi–thousand watt FM station is broadcasting near a 50 mW wireless microphone, even a well-suppressed second harmonic can have an RF field strength comparable to the microphone or intercom signal because the second harmonic of FM 88 MHz is 176 MHz, which is in the middle of television Channel 7. The second harmonic of FM 107 MHz is 214 MHz, which is in the middle of Channel 13. Thus, if a VHF wireless system is to be utilized fully, especially with several microphones or intercoms on adjacent frequencies, the wireless receiver must have a very selective front end.

One television channel occupies a 6 MHz wide segment of the VHF band. Channel 7, for example, covers from 174–180 MHz. A wireless intercom occupies about 0.2 MHz (200 kHz). By FCC Part 74 allocation, up to 24 discrete VHF high-band microphones and/or intercoms can be operated in the space of a single television channel. In order to use multiple systems on adjacent frequencies, the wireless microphone/intercom receivers must be very selective and have an excellent capture ratio (see Section 16.10.1.3). On a practical basis, this means using narrow-deviation FM (approximately 12 kHz modulation). Wide-deviation systems (75 kHz modulation or more) can easily cause interference on adjacent microphone/intercom frequencies; such systems also require wide bandwidth receivers that are more apt to be plagued by interference from adjacent frequencies. The trade-off between narrowband FM and wideband FM favor wideband for better overall frequency response, lower distortion, and inherently better SNR versus maximum possible channels within an unused TV channel for equal freedom from interference (max. 6). Poorly designed FM receivers, are also subject to *desensing*. Desensing means the wireless microphone/intercom receiver is muted because another microphone, intercom, television station, or FM station (second harmonic) is transmitting in close proximity; this limits the effective range of the microphone or intercom.

The UHF band equipment is the band of choice and is the only one used by manufacturers today. The wavelength is less than 3 ft (1 m) so the antennas are only 9 in (23 cm). The range is not as good as VHF, because it can sneak through small openings and can reflect off surfaces more readily.

All of the professional systems now are in the following UHF bands:

- A band 710–722 MHz.

- B band of 722–734 MHz
- 728.125–740.500 MHz band.

The FCC has assigned most of the DTV channels between channel 2 and 51, and only four channels between 64 and 69, which is where most of the professional wireless microphones operate.

16.10.1.2 Adjustment of the System's Operating Frequency

Many of the professional wireless microphones are capable of being tuned to many frequencies. In the past the systems were fixed frequency, often because that was the only way they could be made stable. With PLL-synthesized channels (Phase Lock Loop), it is not uncommon for systems to be switch tunable to 100 different frequencies in the UHF band and have a frequency stability of 0.005%. This is especially important with DTV coming into the scene.

16.10.1.3 Capture Ratio and Muting

Capture ratio and muting specifications of the receiver are important. The capture ratio is the ability of the receiver to discriminate between two transmitters transmitting on the same frequency. When the signal is frequency modulated (FM), the stronger signal controls what the receiver receives. The capture ratio is the difference in the signal strength between the capturing transmitter and the captured transmitter that is blanketed. The lower the number, the better the receiver is at capturing the signal. For instance, a receiver with a capture ratio of 2 dB will capture a signal that is only 2 dB stronger than the second signal.

Most systems have a muting circuit that squelches the system if no RF signal is present. To open the circuit, the transmitter sends a special signal on its carrier that breaks the squelch and passes the audio signal.

16.10.1.4 RF Power Output and Receiver Sensitivity

The maximum legal RF power output of a VHF high-band microphone or intercom transmitter is 50 mW; most deliver from 25–50 mW. Up to 120 mW is permissible in the business band (for wireless intercoms) under FCC part 90.217, but even this represents less than 4 dB more than 50 mW. The FCC does not permit the use of high-gain transmitter antennas, and even if they did, such antennas are large and directional so they would not be practical for someone who is moving around. Incidentally, high-gain receiving antennas are also a bad idea because the transmitter is constantly moving around with the performer so much of the received radio signal is actually caught on the bounce from walls, props, and so on (see Section 16.9.2).

Even if an offstage antenna is aimed at the performer, it probably would be aiming at the wrong target. Diversity receiving antenna systems, where two or more antennas pick up and combine signals to feed the receiver, will reduce dropouts or fades for fixed receiver installations.

The received signal level can't be boosted, given the restrictions on antenna and transmitted power, so usable range relies heavily on receiver sensitivity and selectivity (i.e., capture ratio and SNR) as well as on the audio dynamic range. In the pre-1980 time frame, most wireless microphones and intercoms used a simple compressor to avoid transmitter overmodulation. Today, systems include compandor circuitry for 15–30 dB better audio SNR without changing the RF SNR (see Section 16.10.3). This is achieved by building a full-range compressor into the microphone or intercom transmitter, and then providing complementary expansion of the audio signal at the receiver—much like the encoder of a tape noise-reduction system. The compression keeps loud sounds from overmodulating the transmitter and keeps quiet sounds above the hiss and static. The expander restores the loud sounds after reception and further reduces any low-level hiss or static. Companding the audio signal can provide from 80–85 dB of dynamic range compared to the 50–60 dB of a straight noncompanded transmit/receive system using the same deviation.

16.10.1.5 Frequency Response, Dynamic Range, and Distortion

No wireless microphone will provide flat response from 20 Hz–20 kHz, nor is it really needed. Wireless or not, by the time the audience hears the broadcast, film, or concert, the frequency response has probably been reduced to a bandwidth from 40 Hz–15 kHz. Probably the best criteria for judging a handheld wireless microphone system is to compare it to the microphone capsule's naked response. If the transmit/receive bandwidth basically includes the capsule's bandwidth, it is enough. Generally speaking, a good wireless microphone should sound the same as a hard-wired microphone that uses the same capsule. Wireless intercom systems, because they are primarily for speech communication, are less critical with regard to audio bandwidth; 300 Hz–3 kHz is telephone quality, and 50 Hz–8 kHz is excellent for an intercom.

Dynamic range is probably the most critical aspect of performance for natural sound. A good compandor system will provide 80–85 dB of dynamic range, assuming the microphone is adjusted to 100% modulation on the loudest sounds. Leaving a margin of safety by turning down the microphone modulation level sacrifices SNR. Even with extra headroom and a working SNR of 75 dB, the microphone will still have about twice the dynamic range of a typical optical film sound track or television show.

The system should provide at least 40–50 dB SNR with a 10 µV signal and 70–80 dB SNR with an 80 µV signal. This shows up when no audio signal is being transmitted.

When an electret condenser microphone is used, a major limitation in dynamic range can be the capsule itself, not the wireless system. Typically, an electret powered by a 1.5 V battery is limited to about 105 dB *SPL*. Powered by a 9 V battery, the same microphone may be usable to 120 dB *SPL*. The wireless microphone system should be able to provide a high enough bias voltage to ensure adequate dynamic range from the microphone capsule. Although the condenser may be hotter in output level than a dynamic microphone, its background noise level is disproportionately higher, so the overall SNR specification may be lower.

Wireless intercom systems do not need the same dynamic range as a microphone. They do not have to convey a natural musical performance. However, natural dynamics are less fatiguing than highly compressed audio, especially given a long work shift. So aside from greater range, there are other benefits to seeking good SNR in the intercom: 40 dB or 50 dB would be usable, and 60 dB or 70 dB is excellent. An exception is in a very high-noise industrial environment, where a compressed loud intercom is necessary to overcome background noise. A good intercom headset should double as a hearing protector and exclude much of the ambient noise.

Distortion is higher in a wireless system than in a hard-wired system—a radio link will never be as clean as a straight piece of wire. Still, total harmonic distortion (THD) specifications of less than 1% overall distortion are available in today's better wireless microphones. In these microphones, one of the largest contributors to harmonic distortion is the compandor, so distortion is traded for SNR. The wireless intercom can tolerate more THD, but lower distortion will prevent fatigue and improve communication.

16.10.2 Receiving Antenna Systems

RF signal dropout or multipath cancellation is caused by the RF signal reflecting off a surface and reaching a single receiver antenna 180° out-of-phase with the direct signal, Fig. 16-145. The signal can be reflected off surfaces such as armored concrete walls, metal grids, vehicles, buildings, trees, and even people.

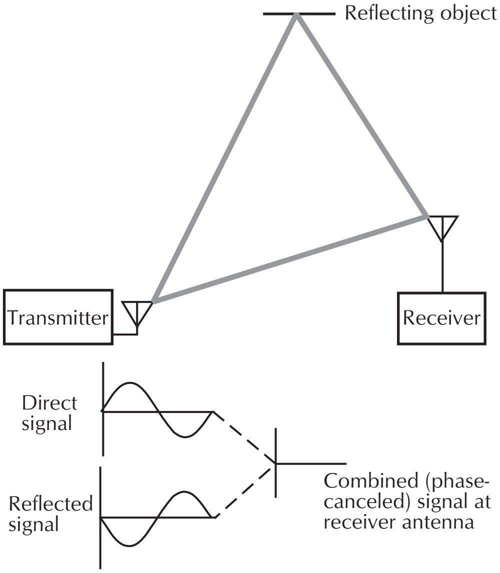


Figure 16-145. Phase cancellation of radiofrequency signals due to reflections.

Although you can often eliminate the problem by experimenting with receiver antenna location, a more foolproof approach is to use a space diversity system where two or more antennas pick up the transmitted signal, as shown in Fig. 16-146. It is highly unlikely that the obstruction or multipath interference will affect two or more receiver antennas simultaneously.

There are three diversity schemes: *switching diversity, true diversity*, and *antenna combination.*

- **Switching Diversity.** In the switching diversity system, the RF signals from two antennas are compared, and only the stronger one is selected and fed to one receiver.
- **True Diversity.** This receiving technique uses two receivers and two antennas set up at different positions, Fig. 16-147. Both receivers operate on the same frequency. The AF signal is taken from the output of the receiver that at any given moment has the stronger signal at its antenna. The probability of no signal at both antennas at the same time is

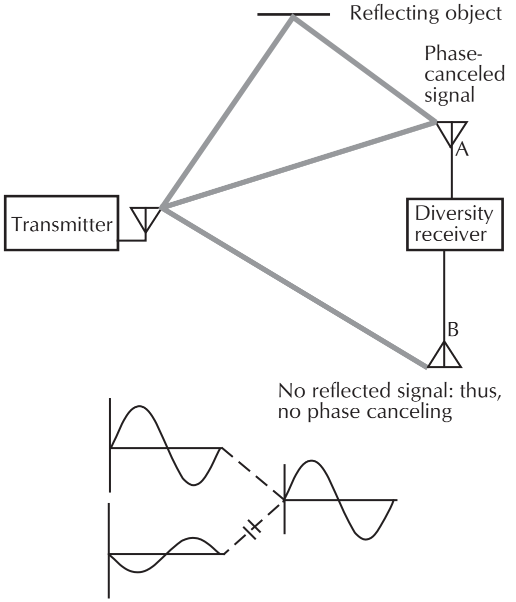


Figure 16-146. Diversity antenna system used to reduce multipath radiofrequency phase cancellation.

extremely small. The advantages of diversity compared to conventional RF transmission are shown in Fig. 16-148. Only the receiving chain with the better input signal delivers audio output. Not only does this system provide redundancy of the receiving end, but it also combines signal strength, polarity and space diversity.

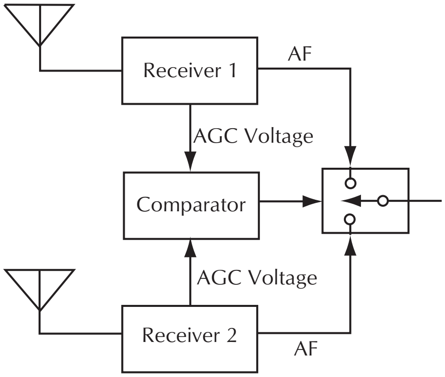


Figure 16-147. Functional diagram of a true diversity receiver.

- **Antenna Combination Diversity.** The antenna combination diversity system is a compromise of the other methods. This system uses two or more antennas, each connected to a wideband RF amplifier to boost the received signal. The signals from both receiving antennas are then actively combined and fed to one standard receiver per microphone. In this way, the receiver always gets the benefit of the signals present at all antennas. There is no switching noise, no change in background noise, and only requires one receiver for each channel. A drawback is the possibility of complete signal cancellation when phase and amplitude relationships due to multipath provide the proper unfavorable conditions.

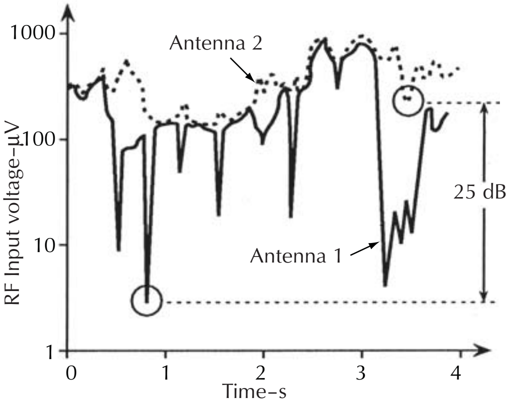


Figure 16-148. Effect of switch-over diversity operation. Solid line RF level at antenna 1 and the dotted line RF level at antenna 2. Courtesy Sennheiser Electronic Corporation.

16.10.2.1 Antenna Placement

It is often common to use a near antenna and a far antenna. The near antenna, which is the one nearest the transmitter, produces the majority of the signal most of the time; in fact, it may even be amplified with an in-line amplifier. The far-field antenna may be one or more antennas usually offset in elevation and position; therefore, the possibility of dropout is greatly reduced. Because the antennas are common to all receivers, many wireless microphones can be used at the same time on the same antenna system. This means that there are fewer antennas and a greater possibility of proper antenna placement.

The following will generally prevent dead spots:

- Do not set up antennas in niches or doorways.
- Keep the antennas away from metal objects including armored concrete walls. Minimum distance: 3 ft (1 m).
- Position the antennas as close as possible to the point where the action takes place.
- Keep antenna cables short to keep RF losses at a minimum. It is better to use longer AF leads instead.

Note: If long runs of antenna cable are used, be sure they are of the low-loss type.

- Make a walkaround test, i.e., operate the transmitter at all positions where it will be used later. Mark all points where field strength is weak. Try to improve reception from these points by changing the antenna position. Repeat this procedure until the optimum result is achieved.

Interference is mainly caused by spurious signals arriving at the receiver input on the working frequency. These spurious signals may have various causes:

- Two transmitters operating on the same frequency (not permissible).
- Intermodulation products of a multichannel system whose frequencies have not been selected carefully enough.
- Excessive spurious radiation from other radio installations—e.g., taxi, police, CB-radio, etc.
- Insufficient interference suppression on electric machinery, vehicle ignition noise, etc.
- Spurious radiation from electronic equipment—e.g., light control equipment, digital displays, synthesizers, digital delays, computers, etc.

16.10.3 Companding

Two of the biggest problems with using wireless microphones are SNR and dynamic range. To overcome these problems, the signal is compressed at the transmitter and expanded at the receiver. Figs. 16-144 and 16-149 graphically illustrate how and what this can accomplish with respect to improving the SNR and reducing the susceptibility to low-level incidental FM modulation, such as buzz zones.

As the typical input level changes by a factor of 80 dB, the audio output to the modulator undergoes a contoured compression, so a change in input audio level is translated into a pseudologarithmic output. This increases the average modulation level, which reduces all forms of interference encountered in the transmission medium.

By employing standard narrowband techniques at the receiver, the recovered audio is virtually free of adjacent channel and spurious response interference. In addition, up to ten times the number of systems can be operated simultaneously without cross-channel interference. The ability of the receiver to reject all forms of interference is imperative when utilizing expansion and compression techniques because the receiver must complimentarily expand the audio component to restore the original signal integrity.

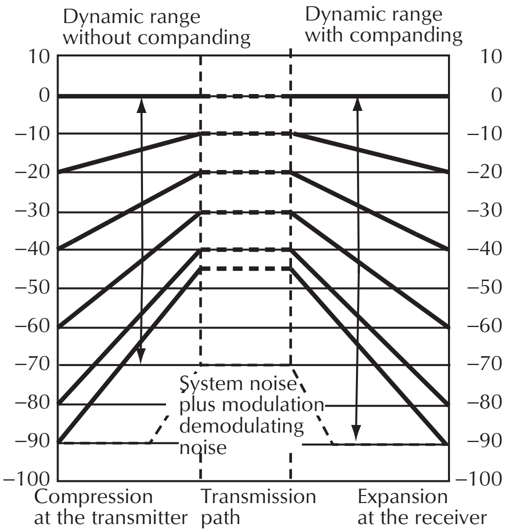


Figure 16-149. Compression and expansion of the audio signal. Notice the −80 dB signal is not altered, and the −20 dB signal is altered significantly.

16.10.4 Waterproof Wireless Microphone Systems

Wireless microphones that are worn are very useful for coaching all forms of athletics including swimming and aquatic aerobics. If the instructor always stays on the pool deck, a weatherproof system might be adequate. If the instructor is in the water, a completely submersible and waterproof system will be required.

Hydrophonics assembles a completely waterproof and submersible wireless microphone system. Assembled with Telex components, the system includes a headset microphone with a special waterproof connector and a Telex VB12 waterproof beltpack transmitter. The transmitter can operate on a 9 V alkaline battery or a 9 V NiMH rechargeable battery. The rechargeable battery is recommended as it does not require removing the battery from the transmitter for recharging and therefore reduces the chance of water leaking into the transmitter housing. The receiver is a Telex VR12 for out-of-pool operation, and can be connected to any sound system the same way as any other wireless microphone.

An interesting thing about this system is you can dive into the water while wearing the system and come up and immediately talk as the water drains out of the windscreen rapidly.

The DPA Type 8011 hydrophone, Fig. 16-150, is a 48 V phantom powered waterproof microphone specially designed to handle the high sound pressure levels and the high static ambient pressure in water and other fluids. The hydrophone uses a piezoelectric sensing element, which is frequency compensated to match the special acoustic conditions under water. A 10 m high-quality audio cable is vulcanized to the body of the hydrophone and fitted with a standard three-pin XLR connector. The output is electronically balanced and offers more than 100 dB dynamic range. The 8011 hydrophone is a good choice for professional sound recordings in water or under other extreme conditions where conventional microphones would be adversely affected.

Figure 16-150. DPA 8011 hydrophone. Courtesy DPA Microphones A/S.

16.11 Multichannel Wireless Microphone and Monitoring Systems

By Joe Ciaudelli and Volker Schmitt

16.11.1 Introduction

The use of wireless microphones and monitoring systems has proliferated in the past few years. This is due to advancements in technology, a trend towards greater mobility on stage, and the desire to control volume and equalization of individual performers. Consequently, installations in which a number of wireless microphones, referred to as channels, are being used simultaneously, have increased dramatically. Now theatres and studios with large multichannel systems, greater than thirty channels, are common. Systems of this magnitude are a difficult engineering challenge. Careful planning, installation, operation, and maintenance are required.

Wireless systems require a transmitter, transmit antenna, and receiver to process sound via radio frequency (RF) transmission. First, the transmitter processes the signal and superimposes it on a carrier through a process called *modulation*. The transmit antenna then acts as a launch pad for the modulated carrier and broadcasts the signal over the transfer medium: air. The signal must then travel a certain space or distance to reach the pickup element, which is the receiving antenna. Finishing up the process, the receiver —which selects the desired carrier—strips off the signal through demodulation, processes it, and finally reconstitutes the original signal. Each wireless channel needs to operate on a unique frequency.

16.11.2 Frequencies

Manufacturers generally produce wireless microphones on ultrahigh frequencies (UHF) within the TV band with specifications outlined by government agencies such as the Federal Communications Commission (FCC). The wavelength is inversely proportional to the frequency. Higher frequencies have shorter wavelengths. UHF frequencies (450–960 MHz) have a wavelength of less than one meter. They have excellent reflective characteristics. They can travel through a long corridor, bouncing off the walls, losing very little energy. They also require less power to transmit the same distance compared to much higher frequencies, such as microwaves. These excellent wave propagation characteristics and low power requirements make UHF ideal for performance applications.

16.11.3 Spacing

In order to have a defined channel, without crosstalk, a minimum spacing of 300 KHz between carrier frequencies should be employed. A wider spacing is even more preferable since many receivers often exhibit desensitized input stages in the presence of closely spaced signals. However, caution should be used when linking receivers with widely spaced frequencies to a common set of antennae. The frequencies need to be within the bandwidth of the antennas.

16.11.4 Frequency Deviation

The modulation of the carrier frequency in an FM system greatly influences its audio quality. The greater the deviation, the better the high-frequency response and the dynamic range. The trade-off is that fewer channels can be used within a frequency range. However, since audio quality is usually the priority, wide deviation is most desirable.

16.11.5 Frequency Coordination

Multichannel wireless microphone systems can be especially difficult to operate, as they present several special conditions. Multiple transmitters moving around a stage will result in wide variations of field strength seen at the receiver antenna system. This makes frequency selection to avoid interference from intermodulation (IM) products highly critical. This is even more challenging in a touring application since the RF conditions vary from venue to venue. In this case, the mix of frequencies is constantly changing. The daunting task to coax each of these variables to execute clear audio transmission can be achieved through careful frequency coordination.

Intermodulation is the result of two or more signals mixing together, producing harmonic distortion. It is a common misconception that intermodulation is produced by the carrier frequencies mixing within the air. Intermodulation occurs within active components, such as transistors, exposed to strong RF input signals. When two or more signals exceed a certain threshold, they drive the active component into a non-linear operating mode and intermodulation (IM) products are generated. This usually happens in the RF section of the receiver, in antenna amplifiers, or the output amplifier of a transmitter. In multichannel operation, when several RF input signals exceed a certain level the intermodulation products grow very quickly. There are different levels of intermodulations defined by the number of addition terms.

In any wireless system with three or more frequencies operating in the same range, frequency coordination is strongly advised.

It is necessary to consider possible IM frequencies that might cause problems for the audio transmission. The 3rd and 5th harmonics, in particular, might raise interference issues.

The following signals may be present at the output of a nonlinear stage:

Fundamentals:	F1 and F2
Second Order:	2F1, 2F2, F1±F2, F2–F1
Third Order:	3F1, 3F2, 2F1±F2, 2F2±F1
Fourth Order:	4F1, 4F2, 2F1±2F2, 2F2±2F1
Fifth Order:	5F1, 5F2, 3F1±2F2, 3F2±2F1
Additional higher orders	

As a result, the intermodulation frequencies should not be used, as those frequencies are virtual transmitters. The fundamental rule *never use two transmitters on the same frequency* is valid in this case. However, even-order products are far removed from the fundamental frequencies and, for simplicity, are therefore omitted from further considerations. Signal amplitude rapidly diminishes with higher-order IM products, and with contemporary equipment design, consideration of IM-products can be limited to 3rd and 5th order only.

For multichannel applications such as those on Broadway (i.e., 30+ channels), the intermodulation products can increase significantly and the calculation of intermodulation-free frequencies can be done by special software. By looking only at the third harmonic distortion in a multichannel system, the number of third-order IM products generated by multiple channels is:

- 2 channels result in 2.
- 3 channels result in 9.
- 4 channels result in 24.
- 5 channels result in 50.
- 6 channels result in 90.
- 7 channels result in 147.
- 8 channels result in 225.
- ….
- 32 channels result in 15,872 third-order IMproducts.

Adding more wireless links to the system will increase the number of possible combinations with interference potential logarithmically: n channels will result in ($n3 - n2$)/2 third-order IM-products. Equal frequency spacing between RF carrier frequencies inevitably results in two- and three-signal intermodulation products and must be avoided!

The RF level and the proximity define the level of the intermodulation product. If two transmitters are close, the possibility of intermodulation will increase significantly. As soon as the distance between two transmitters is increased, the resulting intermodulation product decreases significantly. By taking this into consideration, the physical distance between two or more transmitters is important. If a performer needs to wear two bodypack transmitters, it is recommended to use two different frequency ranges and to wear one with the antenna pointing up and the other with it pointing down.

If the number of wireless channels increases, the required RF bandwidth increases significantly, Fig. 16-151.

External disturbing sources such as TV transmitters, taxi services, police services, digital equipment, etc., also have to be taken into consideration. Fortunately, the

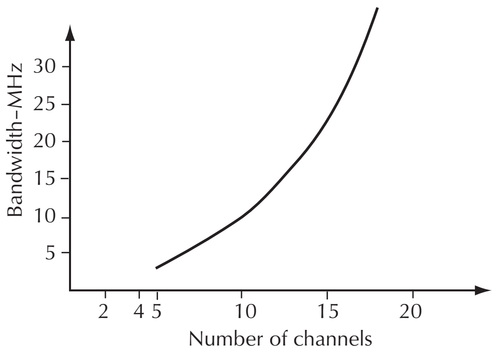


Figure 16-151. Bandwidth required for multi-channel systems.

screening effect of buildings is rather high (30–40 dB). For indoor applications, this effect keeps strong outside signals at low levels. A significant problem can occur when poorly screened digital equipment is working in the same room. These wideband disturbing sources are able to interfere with wireless audio equipment. The only solution to this problem is to replace the poorly screened piece of equipment with a better one.

Other RF-systems that have to be considered for compatibility are:

1. TV stations "On-Air."
2. Wireless intercoms.
3. IFBs.
4. Wireless monitor systems.
5. Other wireless systems.

Compatibility between components of a system is achieved if the following requirements are met: each link in a multichannel wireless system functions equally well with all other links active and no single link—or any combination of multiple links—causes any interference.

If the transmitter of a wireless mic channel is switched off, its complementary receiver should also be switched off or muted at the mixing console. A receiver that does not see its transmitter will try to latch onto a nearby signal. That signal may be an intermodulation product. The receiver will then try to demodulate this signal and apply it to the speaker system.

Equipment can be designed to minimize intermodulation. A specification known as *intermodulation rejection* or *suppression* is a measure of the RF input threshold before intermodulation occurs. For a well designed receiver, this specification will be 60 dB or greater. An intermodulation rejection of 60 dB means that intermodulation products are generated at input levels of approximately 1 mV. The highest quality multichannel receivers currently available feature an intermodulation rejection of >80 dB. If high-quality components are used, having an intermodulation suppression of 60 dB or greater, only the third-order products need to be considered.

16.11.6 Transmitter Considerations

Transmitters are widely available as portable devices, such as handheld microphones, bodypacks, and plug-on transmitters and are produced in stationary form as stereo monitors. When transmitting signals for most wireless applications via air, FM modulation is generally used; in doing so, one must also improve the sound quality in a variety of ways.

An RF transmitter works like a miniature FM radio station. First, the audio signal of a microphone is subjected to some processing. Then the processed signal modulates an oscillator, from which the carrier frequency is derived. The modulated carrier is radiated via the transmitter's antenna. This signal is picked up by a complementary receiver via its antenna system and is demodulated and processed back to the original audio signal.

16.11.6.1 Range and RF Power

Transmitter power is a rating of its potential RF signal strength. This specification is measured at the antenna output. The range of a wireless transmission depends on several factors. RF power, the operating frequency, the setup of the transmitter and receiver antennas, environmental conditions, and how the transmitter is held or worn are all aspects that determine the overall coverage of the system. Therefore, power specifications are of only limited use in assessing a transmitter's range, considering these variable conditions. Also, battery life is associated with RF output power. Increased power will reduce battery life with only a moderate increase in range.

Using RF wireless microphone transmitters with the right amount of RF output power is important to ensure total system reliability. There is a common misconception that higher power is better. However, in many applications high power can aggravate intermodulation (IM) distortion, resulting in audible noises.

First of all, the applied RF output power must fall within the limit allowed by each country's legislation. In the United States, the maximum RF output power for wireless microphones is limited to 250 mW. In most of the countries in Europe this figure is 50 mW, while in Japan it is only 10 mW. Despite the 10 mW limitation, many multichannel wireless microphones are operating in Japan. This is achieved by careful attention to factors like antenna position, use of low loss RF cables and RF gain structure of the antenna distribution system.

There are indeed some applications in which more RF output power is an appropriate measure; a perfect example would be a golf tournament, as the wireless system needs to cover a wide area. There are usually only a few wireless microphones in use at this type of function, and those microphones are generally not in close proximity to each other.

If transmitters with high RF power are close together, intermodulation usually occurs. At the same time, the RF noise floor in the performance area is increased. As a matter of fact, a transmitter in close proximity to another transmitter will not only transmit its own signal, but it will also receive the signal and add this to the RF amplifier stage.

16.11.6.2 Dc-to-Dc Converter

Transmitters should be designed to provide constant RF output power and frequency deviation throughout the event being staged. This can he achieved through the use of a dc-to-dc converter circuit. Such a circuit takes the decaying battery voltage as its input and regulates it to have a constant voltage output. Once the voltage of the batteries drops below a minimum level, the dc-to-dc converter shuts off, almost instantaneously. The result is a transmitter that is essentially either off or on. While it is on, the RF output power, frequency deviation, and other relevant specifications remain the same. Transmitters without regulation circuits, once the battery voltage begins to drop, will experience reduced range and audio quality.

16.11.6.3 Audio Processing

To improve the audio quality, several measures are necessary because of the inherent noise of the RF link.

16.11.6.3.1 Pre- and De-Emphasis

This method is a static measure that is used in most of the FM transmissions. By increasing the level of the higher audio frequencies on the transmitter side, the signal-to-noise ratio is improved because the desired signal is above the inherent noise floor of the RF link.

16.11.6.3.2 Companding

The compander is a synonym for *compressor* on the transmitter side and for *expander* on the receiving end. The compressor raises low audio level above the RF noise floor. The expander does the mirror opposite and restores the audio signal. This measure increases the signal-to-noise ratio to CD quality level.

16.11.6.3.3 Spurious Emissions

Apart from the wanted carrier frequency, transmitters can also radiate some unwanted frequencies known as spurious emissions. For large multichannel systems these spurious frequencies cannot be ignored. They can be significantly reduced through elaborate filtering and contained by using a well-constructed, RF tight metal housing for the transmitter. Also, an RF tight transmitter is less susceptible to outside interference.

A metal housing is important not only for its shielding properties, but also its durability. These devices usually experience much more abuse by actors and other talent than anyone ever predicts.

16.11.6.4 Transmitter Antenna

Every wireless transmitter is equipped with an antenna, which is critically important to the performance of the wireless system. If this transmitter antenna comes in contact with the human body, the transmitted wireless energy is reduced and may cause audible noises known as *drop-outs*. This effect of detuning the antenna on contact is called body absorption.

For this reason, talent should not touch the antenna while using handheld microphones. Unfortunately, there is no guarantee that they will follow this recommendation. Taking this into account, optimized antenna setup at the receiver side and the overall RF gain structure of the system becomes critical.

This same effect can occur when using bodypack transmitters, especially if the talent is sweating. A sweaty shirt can act as a good conductive material to the skin. If the transmitter antenna touches it, reduced power and thus poor signal quality may result. In this case, a possible approach is to wear the bodypack upside down near or attached to the belt, with the antenna pointing down. Sometimes this measure does not work because the talent will sit on the antenna. In this case, a possible solution is keeping the transmitter in the normal position and fitting a thick-walled plastic tube over the antenna, such as the type that is used for aquarium filters.

16.11.7 Receiver Considerations

The receiver is a crucial component of wireless audio systems, as it is used to pick the desired signal and transfer its electrical information into an audio signal. Understanding basic receiver design, audio processing, squelch, and diversity operation can help ensure optimum performance of the system.

Virtually all modern receivers feature superheterodyne architecture, in which the desired carrier is filtered out from the multitude of signals picked up by the antenna, then amplified and mixed with a local oscillator frequency to generate the difference: *intermediate frequency*. This IF undergoes more controlled discrimination and amplification before the signal is demodulated and processed to restore the output with all the characteristics and qualities of the original.

Audio signal processing of a receiver is the mirror opposite of the transmitter. Processing done in the transmitters often include pre-emphasis (boosting high audio frequencies) as well as compression. These are reversed in the receiver by the de-emphasis and the expander circuit.

An inherent RF noise floor exists in the air. The squelch setting should be set above this noise level. This acts as a noise gate that mutes the audio output if the wanted RF signal falls below a threshold level. This prevents a blast of white noise through the PA if the RF signal is completely lost. If the squelch setting is too low, the receiver might pick the noise floor and this noise can be heard. If the squelch setting is too high the range of the wireless microphone is reduced.

16.11.7.1 RF Signal Level

Varying RF signal strength is mainly due to multi-path propagation, absorption and shadowing. These are familiar difficulties also experienced with car radios in cities.

Audible effects due to low RF signals, known as *dropouts*, can occur even at close range to the receiver due to multipath propagation. Some of the transmitted waves find a direct path to the receiver antenna and others are deflected off a wall or other object. The antenna detects the vector sum, magnitude and phase, of direct and deflected waves it receives at any particular instant. A deflected wave can diminish a direct wave if it has different phase, resulting in an overall low signal. This difference in phase is due to the longer path a deflected wave travels between the transmitter and receiver antennae and any phase reversal occurring when it hits an object. This phenomenon needs to be addressed in an indoor application since the field strength variation inside a building with reflecting walls is 40 dB or more. It is less critical outside.

RF energy can be absorbed by nonmetallic objects resulting in low signal strength. As stated previously, the human body absorbs RF energy quite well. It is important to place antennas correctly to minimize this effect.

Shadowing occurs when a wave is blocked by a large obstacle between the transmitter and receiver antennas. This effect can be minimized by keeping the antennas high and distance of ½ wavelength away from any large or metallic objects.

These problems are addressed by a diversity receiver. A diversity system is recommended even if only one channel is in operation. Large multichannel systems are only possible with diversity operation.

There are different kinds of diversity concepts available. Antenna switching diversity uses two antennas and a single receiving circuit. If the level at one antenna falls below a certain threshold it switches to the other antenna. This is an economical architecture but it leaves the chance that the second antenna could be experiencing an even lower signal then the one that falls below the threshold level. Another approach is the switching of the audio signal of two independent receiver units where each receiver unit is connected to its own antenna. This is known as *true diversity*. This technique improves the effective RF receiving level by greater than 20 dB. Depending on the diversity concept, active switching between the two antennas is a desired result.

The minimum distance between the two diversity antennas is very often an issue of debate. A minimum of ¼ of a wavelength of the frequency wave seems to be a good approach. Depending on the frequency, 5–6 inches is the minimum distance. In general, a greater distance is preferred.

16.11.8 Antennas

The position of the antenna and the correct use of its related components—such as the RF cable, antenna boosters, antenna attenuators, and antenna distribution systems—are the key to trouble-free wireless transmission. The antennas act as the eyes of the receiver, so the best results can be achieved by forming a direct line of sight between the transmitter antenna and receiver antenna of the system.

Receiving and transmitting antennas are available as omnidirectional and directional variants.

For receiving, omnidirectional antennas are often recommended for indoor use because the RF signal is reflected off of the walls and ceiling. When working outside, one should choose a directional antenna since there are usually little to no reflections outdoors, and this directivity will help to stabilize the signal. In general, it is wise to keep an antenna toolbox that contains both omnidirectional and directional antennas for use in critical RF situations, since they transmit and receive signals differently.

Omnidirectional antennas transmit or receive the signal by providing uniform radiation or response only in one reference plane, which is usually the horizontal one parallel to the earth's surface. The omni-directional antenna has no preferred direction and cannot differentiate between a wanted and an unwanted signal.

If a directional antenna is used, it will transmit or receive the signal in the path it is pointing toward. The most common types are the yagi antenna and the log-periodic antenna, which are often wide-range frequency antennas covering the whole UHF range. In an outdoor venue, the desired signal can be received and the unwanted signal from a TV station can be rejected to a certain degree by choosing the correct antenna position. A directional antenna also transmits or receives only in one plane, like an omnidirectional antenna.

Several types of omnidirectional and directional antennas also exist for specific conditions. The telescopic antenna is an omnidirectional antenna and often achieves a wide range (450–960 MHz). If telescopic antennas are in use they should be placed within the line of sight of the counterpart antenna. They should not, for example, be mounted inside a metal flight case with closed doors as this will reduce the RF field strength from the transmitter and compromise the audio quality.

System performance will be raised considerably when remote antennas are used. A remote antenna is one that is separated from the receiver or transmitter unit. These antennas can be placed on a stand such as that for a microphone. This will improve the RF performance significantly. However, when using remote antennas, some basic rules need to be considered. Once again, a clear line of sight should be established between the transmitter and receiver antenna, Fig. 16-152.

Placing antennas above the talent increases the possibility the transmitter and receiver remain within line of sight, ensuring trouble-free transmission.

If a directional antenna is used, the position of the antenna and the distance to the stage is important. One common setup is pointing both receiving antennas toward the center of the stage. Once again, a line of sight between the receiver and transmitter antennas is best for optimum transmission quality.

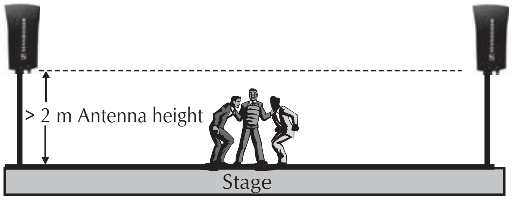


Figure 16-152. Placing antennas above the talent increases the possibility the transmitter and receiver remain within line of sight, ensuring trouble-free transmission.

Directional and omnidirectional antennas do have a preferred plane, which is either the horizontal or vertical plane. If the polarization between the transmitter and receiver antenna is different, this will cause some significant loss of the RF level. Unfortunately, it is not possible to have the same polarization of the antennas all of the time. In a theatrical application, the antenna is in a vertical position when the actress or actor walks on the stage. The polarization of the transmitter may change to the horizontal position if a scene requires the talent to lie down or crawl across the stage. In this case, circular polarized antennas can help. These kinds of antennas can receive the RF signal in all planes with the same efficiency.

Because the polarization of the antenna is critical and telescopic antennas are often used, it is not recommended to use the receiver antennas strictly in a horizontal or vertical plane. Rather, angle the antennas slightly as this will minimize the possibility that polarization would be completely opposite between transmitter and receiver.

One last note: The plural form for the type of antenna discussed in this article is *antennas*. Antennae are found on insects and aliens.

16.11.8.1 Antenna Cables and Related Systems

Antenna cables are often an underestimated factor in the design of a wireless system. The designer must choose the best cable for practical application, depending on the cable run and the installation, Table 16-3. As the RF travels down the cable its amplitude is attenuated. The amount of this loss is dependent on the quality of the cable, its length and the RF frequency. The loss increases with longer cable and higher frequencies. Both of these effects must be considered for the design of a wireless microphone system.

RF cables with a better specification regarding RF loss are often thicker. These are highly recommended for fixed installations. In a touring application, in which the cable must be stored away each day, these heavier cables can be very cumbersome.

Table 16-3. Different Types of RF Cables with Various Diameters and the Related Attenuation for Different Frequencies.

Cable Type	Frequency (MHz)	Attenuation (db/100′)	Attenuation (dB/100m)	Cable Diameter (inches/ mm)
RG-174/U	400 700	19.0 27.0	62.3 88.6	0.110 / 2.8
RG-58/U	400 700	9.1 12.8	29.9 42.0	0.195 / 4.95
RG-8X	400 700	6.6 9.1	21.7 29.9	0.242 / 6.1
RG-8/U	400 700	4.2 5.9	13.2 19.4	0.405 / 10.3
RG-213	400 700	4.5 6.5	14.8 21.8	0.405 / 10.3
Belden 9913	400 700	2.7 3.6	8.9 11.8	0.405 / 10.3
Belden 9913F 9914	400 700	2.9 3.9	9.5 12.8	0.405 / 10.3

Source: Belden Master Catalogue.

As any RF cable has some RF attenuation, cable length should be as short as possible without significantly increasing the distance between the transmitter and receiver antennas. This aspect is important for receiving applications but is even more critical for the transmission of a wireless monitor signal.

In a receiving application, it is important to consider losses from the cable as well as from any splitter in the antenna system during the design and concept stage of a wireless microphone system. If the losses in the system are small, an antenna booster should not be used. In this case, any drop-out is not related to the RF loss in the antenna system; instead, it is more often related to the antenna position and how the transmitter is used and worn during the performance. An antenna booster is recommended if the loss in the antenna system is greater than 6 dB.

If an antenna booster is necessary, it should be placed as close as possible to the receiving antenna. Antennas with a built-in booster are known as *active antennas*. Some of these have a built-in filter, only allowing the wanted frequency range to be amplified. This is recommended because it reduces the possibility of intermodulation.

Two antenna boosters should not be used back-to-back when the RF cable run is very long. The second antenna booster would be overloaded by the output of the first amplifier and would produce intermodulation.

Special care must be taken when using an antenna booster if the transmitter comes close to the receiver antenna. The resulting strong signal could drive the antenna booster past its linear operation range, thus producing intermodulation products. It is recommended to design and install a system such that the transmitter remains at least 10 feet from the receiver antenna at all times.

Another important factor is the filter at the input stage of the antenna booster. The approach is to reduce the amount of unwanted signals in the RF domain as much as possible. This is another measure to reduce the possibility of intermodulation of this amplifier.

Also, signals that come from a TV station—such as a digital television (DTV) signal—are unwanted signals and can be the reason for Intermodulation products in the first amplifier.

If the free TV channel between the DTV should be used for wireless microphone transmission, the DTV signals might cause the problems. To reduce the effect of DTV signals, a narrow input filter will help to overcome the possible effect of Intermodulation.

Often a narrower filter at the input stage of a wireless receiver is preferable. This will often work for fixed installations because there are decreased possibilities that the RF environment will change. This is especially the case when the RF environment is difficult and a lot of TV stations or other wireless systems are operating.

16.11.8.2 Splitter Systems

Antenna splitters allow multiple receivers to operate from a single pair of antennas. Active splitters should be used for systems greater than four channels so that the amplifiers can compensate for the splitter loss. Security from interference and intermodulation can be enhanced by filtering before any amplifier stage. As an example, a thirty-two-channel system could be divided into four subgroups of eight channels. The subgroups can be separated from each other by highly selective filters. The subgroups can then be considered independent of each other. In this way, frequency coordination only needs to be performed within each group. It is much easier to coordinate eight frequencies four times than to attempt to coordinate a single set of thirty-two frequencies, Fig. 16-153.

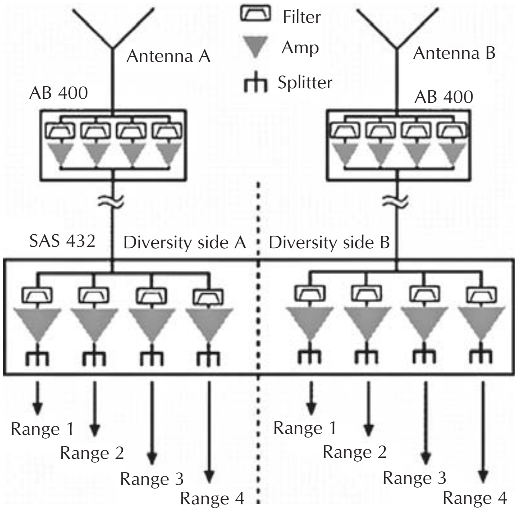


Figure 16-153. Diversity antenna setup with filtered boosters, long antenna cables, and active splitter with selective filtering.

16.11.9 Wireless Monitor Systems

Wireless monitor systems are essential for stage-bound musical productions. Perhaps the biggest advantage of a wireless monitor system is the ability to use an individual monitor mix for each musician on stage. Furthermore, a wireless monitor system significantly reduces the amount of, or even eliminates, monitor speakers in the performance area. This results in lower risk of feedback and a more lightweight, compact monitor system.

Some special precautions must be taken before using wireless monitor systems. In most cases, this signal is a stereo signal. This multiplexed signal is more sensitive to dropouts and static and multipath situations. For long range applications, mono operation can improve system performance.

If wireless microphones and wireless monitor systems are used in parallel, those systems should be separated in a way that the frequencies are at least 8 MHz apart and that the physical distance between the transmitter and the receiver is maximized. This will reduce the risk of blocking—an effect that desensitizes a receiver and prevents the reception of the desired signal. Therefore, if a bodypack wireless mic transmitter and a wireless monitor receiver are both attached to the same talent, those devices should *not* be mounted directly beside each other.

When musicians use the same monitor mix, one transmitter can be used to provide the RF signal to more than one wireless monitor receiver. If individual mixes are desired, each mix requires its own transmitter operating on a unique frequency. To avoid intermodulation disturbances, the wireless monitor transmitters should be combined, and the combined signal should then be transmitted via one antenna. Active combiners are highly recommended. Passive combiners suffer from signal loss and high crosstalk. An active combiner isolates each transmitter by around 40 dB from the other and keeps the RF level the same (0 dB gain), thus minimizing intermodulation. Again, intermodulation is a major issue within the entire wireless concept. When using stereo transmission, it is even more critical.

When considering an external antenna, one important factor must be taken into consideration: the antenna cable should be as short as possible to avoid losses via the RF cable. A directional external antenna is recommended to reduce multipath situations from reflections, and it will have some additional passive gain that will increase the range of the system.

If remote antennas are used for the wireless monitor transmitters as well as wireless mic receivers, those antennas should be separated by at least 10–15 feet. Blocking of the receivers, as discussed above, is then avoided. Furthermore, the antennas should not come in direct contact with the metal of the lighting rig. This will detune the antenna and reduce the effective radiated wireless signal.

16.11.10 System Planning for Multichannel Wireless Systems

When putting together a multi-channel wireless microphone system, several items are essential for trouble-free operation. First, you must understand the environment in which the system will be used.

Location. The location of a venue can be determined by using mapping tools on the internet, such as Google Earth. If you figure out the coordinates of the venue, simply plug this information into the FCC homepage, http://www.fcc.gov/fcc-bin/audio/tvq.html. The result shows all transmitters licensed by the FCC in this area. This information will allow the designer of the wireless system to plan which vacant TV channels can be used for wireless audio devices. If there is a TV transmitter close to the location of the wireless microphone system (<70 miles), this TV channel should generally be avoided. Once one knows which TV channels may be used in the area, the designer can use another software tool that calculates the IM-free frequencies and displays possible setups.

Quantity and Frequency Coordination. Determine how many wireless microphones, wireless monitor systems, intercoms, etc. are required or in use for your job. With the information you gathered from step one, you can begin the system design. You now have the available TV channels and the number of wireless systems you want to use.

With this know-how you can start the frequency coordination of your system inside the vacant TV channels. This is supported by software that is available from various companies. The key here is to prevent intermodulation products (unwanted frequencies generated by harmonic distortion) from interfering with the wanted frequencies of your wireless systems.

A check in the venue is also necessary. If you have the chance, scout the location with a spectrum analyzer, Fig. 16-154. With this tool, you can verify that the information from the internet is correct. Alternately, you can scroll through the tunable frequencies of your wireless receivers to scan the RF activity in the venue. Many receivers also have an auto scan function to find open frequencies. This cross-check is necessary to find out whether other wireless devices are in use that you do not have on your list, which could interfere with your signal during operation.

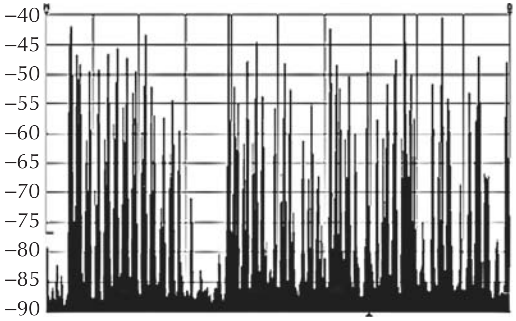


Figure 16-154. Plot of the RF spectrum in Athens outside the Olympic Stadium (450–960 MHz).

Tune Your Components. Set your individual transmitters and corresponding receivers to their coordinated frequencies. Switch on all components and perform a final test of compatibility. Physically space the transmitters a couple feet apart and at least 10 feet from the receiving antenna. Listen for any interference. Compatibility between components of a system is achieved if the following requirements are met: each link in a multichannel wireless system functions equally well with all other links active and no single link—or any combination of multiple links—causes interference.

16.11.11 Future Considerations: Digital Wireless Transmission

Digital is a buzz word that many presume solves all the technical issues we face today. More and more digital equipment, such as mixing consoles, audio signal processors, and the like, are used for several applications, as a digital audio signal chain offers many advantages. A digital signal on a wire (i.e., fiber optic cable) is easier to handle than on a copper wire because 48, 64, or more audio channels can be transported on one thin fiber optic cable. If an audio signal is already in the digital domain, it makes sense to keep it in this domain as long as possible.

As for digital wireless transmission, a digital *wireless* system is beneficial when the sound, occupied RF spectrum, and battery lifetime is as good or even better than an analog system. On top of this, latency (time delay between input and output) is always a very important topic to keep in mind.

16.11.11.1 Starting with Sound and the Related Data Rate

The best sound can be expected if there is no audio data compression used in the wireless system. This will lead to a very high data rate.

- Minimum for 20 kHz audio and approximately 110 dB dynamic range: 18 Bit × 48 kHz = 0.864 Mbit/s.
- Necessary overhead (framing, channel coding) leads to even higher data rate (factor approximately 1.5 [1.296 Mbit/s]).

When transmitting this high amount of data, it is no longer possible to use a simple and robust digital modulation scheme like FSK (frequency shift keying), ASK (amplitude shift keying), or PSK (phase shift keying), because these concepts will be not able to fulfill the spectrum mask, ≤ 200 kHz of occupied RF spectrum, defined by the FCC. Even if this constraint didn't exist, greater occupied RF spectrum could inhibit large multichannel systems.

To improve this, it is necessary to use a more complex modulation scheme with narrow filtering, Fig. 16-155.

The amplitude and the phase of the transmitted signal must be very precise when using this approach. Behind every point of the constellation diagram, a digital word is deposited, which the receiver has to pick up and transfer back into an audio signal.

This requires a very linear RF amplifier. This is a current-hungry device. The unwanted effect is reduced

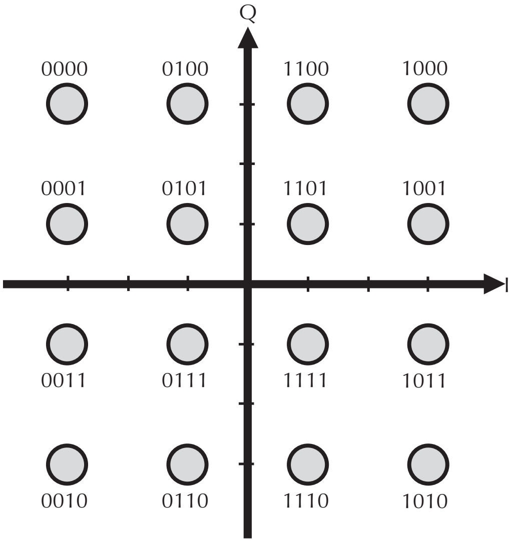


Figure 16-155. Constellation diagram of a 16 QAM modulation.

battery life of transmitters and portable receivers. By driving the RF amplifier with a better efficiency, the occupied RF spectrum will increase in an undesirable manner.

If the data rate described above can be reduced, the modulation scheme can be simplified and the amplified RF can be used in a more efficient way to conserve battery power and increase operational time.

To reduce the amount of digital data a compression algorithm has to be defined. This algorithm will add some latency to the whole data transmission process. Low latency is especially important during a live performance on stage. If the total latency in a PA system, including contributions from digital mixing consoles, effects, etc., is >10 ms, the timing of the band will be thrown off. Furthermore, if streaming video is projected to accommodate a large audience the picture and sound will be out of sync.

New audio data compression algorithms show good performance with a very low latency. However, audio compression would introduce the possibility of audible artifacts (at least with awkward signals).

As technology improves, there will be solutions to the obstacles described above and digital will become available for wireless transmission.

The key questions for a digital system at this time are:

- Is data compression used?
- What RF spectrum is necessary and how will this impact multichannel systems?
- What is the latency of the system?
- What is the battery lifetime?

16.11.12 Conclusion

Large multichannel wireless systems demand excellent planning, especially in the initial phase, and good technical support. Observing all the above-mentioned items, perfect operation of a system can be guaranteed, even under difficult conditions.

16.12 Microphone Accessories

16.12.1 Inline Microphone Processors

The overall sound of a microphone can often benefit from signal processing, and most mixers provide some basic equalization as a tool for customizing the sound of the microphone. Digital mixers provide an even greater set of tools, including parametric EQ, compression, gain management, and other automated functions. Dedicated signal processing for each microphone in a system provides a real advantage for the user, and some manufacturers are offering this sort of custom processing on a per-microphone basis via processors that plug inline with the microphone. These include automatic gain control, automatic feedback control, control for plosives and the proximity effect, and integrated infrared gates that turn the microphone on and off based on the presence of a person near the microphone. These phantom-powered processors allow for targeted solutions to many problems caused by poor microphone technique.

16.12.1.1 Sabine Phantom Mic Rider Pro Series 3

The series 3 Mic Rider includes infrared gates that turn the microphone on and off based on the presence of a person. The heat-sensing IR module is mounted on the gooseneck or is built in on the handheld version. The IR sensor can be adjusted for both time to turn off of 5–15 s and for distances of 3–9 ft, Figs. 16-156 and 16-157.

16.12.1.2 Sabine Phantom Mic Rider Pro Series 2

The series 2 Mic Rider includes the adjustable IR Gate plus three audio processors: automatic gain control, proximity effect control for controlling increased bass due to the proximity effect, and plosive control for

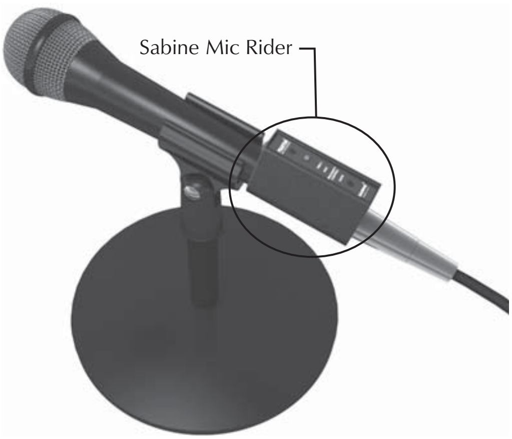


Figure 16-156. The Sabine Inline Mic Rider. Courtesy Sabine, Inc.

Figure 16-157. The Sabine gooseneck Mic Rider with built-in IR sensor. Courtesy Sabine, Inc.

reducing pops and bursts from certain consonant sounds in speech.

16.12.1.3 Sabine Phantom Mic Rider Series 1

The series 1 Mic Rider includes Sabine's patented FBX Feedback Exterminator for maximizing gain before feedback plus the automatic gain control, proximity effect control, and plosive control. A nonadjustable IR gate is also included.

The Phantom Mic Rider works with 48 Vdc phantom power sources that conform to industry standards (DIN standard 45 596 or IEC standard 268-15A). Devices that do not conform can be modified to meet the standard, or external phantom power supplies can be used.

16.12.1.4 Lectrosonics UH400A Plug-On Transmitter

The design of this transmitter was introduced in 1988, in a VHF version aimed at broadcast ENG applications at a time when production crews were being downsized. Converting the popular dynamic microphones of the day to wireless operation eliminated the cable, which was very useful for the two-person production crews that had evolved. During the 20 years that followed, the design continued to evolve to address an ever-increasing variety of applications. The addition of selectable bias voltage allowed the transmitter to power electret microphones. The move to UHF frequencies and a dual-band compandor increased operating range and audio quality. Modifications to the design continued through the present day, leading to the current DSP-based model available in two versions for use with all types of microphones and modest line level signal sources.

The UH400A model has a 12 dB/octave low-frequency roll-off down 3 dB at 70 Hz. The UH400TM model offers an extended low frequency response down 3 dB at 35 Hz, Fig. 16-158. Fig.16-159 is the block diagram of the transmitter.

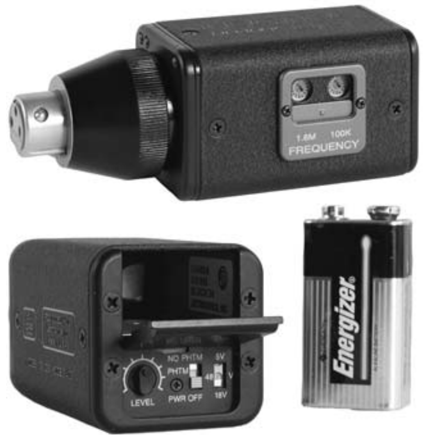


Figure 16-158. Lectrosonics UH400A plug-on transmitter. Courtesy Lectrosonics, Inc.

The most common applications of this transmitter are eliminating the cable between a microphone and a sound or recording system. A prime example is acoustic analysis in a large auditorium or stadium where measurements must be made at multiple locations around the sound system coverage area and extremely long cable runs are not practical. In this case, the wireless not only speeds up the process of making measurements, but it also allows more measurements to be taken, which can improve the final sound system performance.

Digital Hybrid Wireless™ is a patented process that combines a 24 bit digital audio stream with wide deviation FM (U.S. Patent 7,225,135). The process elimi-

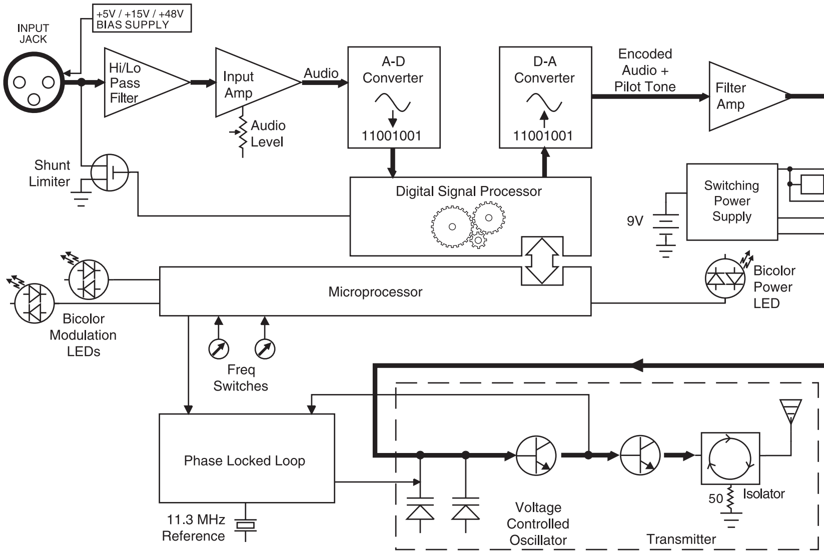

INPUT JACK
+5V / +15V / +48V BIAS SUPPLY
Hi/Lo Pass Filter
Input Amp
Audio
Audio Level
A-D Converter
11001001
D-A Converter
11001001
Encoded Audio + Pilot Tone
Filter Amp
Shunt Limiter
Digital Signal Processor
Switching Power Supply
9V
+3.3v
+3.2v
+1.8v
+6v
-3v
Microprocessor
Bicolor Power LED
Bicolor Modulation LEDs
Freq Switches
Phase Locked Loop
11.3 MHz Reference
Voltage Controlled Oscillator
50
Isolator
Transmitter

Figure 16-159. Block diagram of the Lectrosonics UH400A transmitter. Courtesy Lectrosonics, Inc.

nates a compandor to increase audio quality and expand the applications to test and measurement and musical instrument applications.

Audio is sampled at 88.2 kHz and converted to a 24 bit digital stream. The DSP applies an encoding algorithm that creates what might be likened to an instruction set that is sent to the receiver via an FM carrier. The DSP in the receiver then applies an inverse of the encoding algorithm and regenerates the 24-bit digital audio stream.

An additional benefit of the FM radio link is the ability of the DSP to emulate a compandor for compatibility with analog receivers from Lectrosonics and two other manufacturers.

In the native hybrid mode, the FM deviation is ±75 kHz to provide a wide dynamic range. This wide deviation combined with 100 mW of output power provides a significant improvement in the audio SNR and the suppression of RF noise and interference.

Used with a microphone, the antenna is a dipole formed between the transmitter housing and the microphone body. When plugged into a console or mixer output, the housing of the transmitter is similar to the radiator of a ground plane antenna, with the console or mixer chassis functioning as the ground.

Phantom power can be set to 5, 15,or 48 V or turned off. The transmitter can provide up to 15 mA of current in 5 and 15 V settings, and up to 7 mA in the 48 V setting, allowing it to be used with any type of microphone, including high-end studio condenser models.

The transmitter is available on 9 different frequency blocks in the UHF band between 470 and 692 MHz. Each block provides 256 frequencies in 100 kHz steps.

16.12.1.5 MXL Mic Mate™ USB Adapter

The Mic Mate™, Fig. 16-160, is a USB adapter used to connect a microphone to a Macintosh or PC computer. It uses a 16 bit Delta Sigma A/D converter with $THD + N = 0.01\%$ at sampling rates of 44.1 and 48.0 kHz and includes a three-position analog gain control. It includes a studio-quality USB microphone preamp with a balanced low noise analog input, supplies 48 Vdc phantom power to the microphone, and includes MXL USB Recorder Software for two track recording. There are three different Mic Mates, one for condenser microphones, one for dynamic microphones, and one for news line feeds, video cameras, etc.

MXL
USB Mic Mate

Figure 16-160. MXL Mic Mate USB adapter. Courtesy Marshall Electronics.

16.12.2 Windscreens and Pop Filters

A *windscreen* is a device placed over the exterior of a microphone for the purpose of reducing the effects of breath noise and wind noise when recording out of doors or when panning or gunning a microphone. A windscreen's effectiveness increases with its surface area and the surface characteristics. By creating innumerable miniature turbulences and averaging them over a large area, the sum approaches zero disturbance. It follows that no gain is derived from placing a small foam screen inside a larger blimp type, Fig. 16-161A, whereas a furry cover can bring 20 dB improvement, Fig. 16-161B. Most microphones made today have an integral windscreen/pop filter built in. In very windy conditions, these may not be enough; therefore, an external windscreen must be used.

A. Blimp-type windscreen.

B. Furry cover to surround the windscreen in A.

Figure 16-161. Blimp-type windscreen for an interference tube microphone. Courtesy Sennheiser Electronic Corporation.

With a properly designed windscreen, a reduction of 20–30 dB in wind noise can be expected, depending on the *SPL* at the time, wind velocity, and the frequency of the sound pickup. Windscreens may be used with any type microphone because they vary in their size and shape. Fig. 16-162 shows a windscreen produced by Shure employing a special type polyurethane foam. This material has little effect on the high-frequency response of the microphone because of its porous nature. Standard styrofoam is not satisfactory for windscreen construction because of its homogeneous nature.

A cross-sectional view of a windscreen employing a wire frame covered with nylon crepe for mounting on a 1 inch diameter microphone is shown in Fig. 16-163. The effectiveness of this screen as measured by Dr. V. Brüel of Brüel and Kjaer is given in Fig. 16-164.

Figure 16-162. Shure A81G polyurethane windscreen. Courtesy Shure Incorporated.

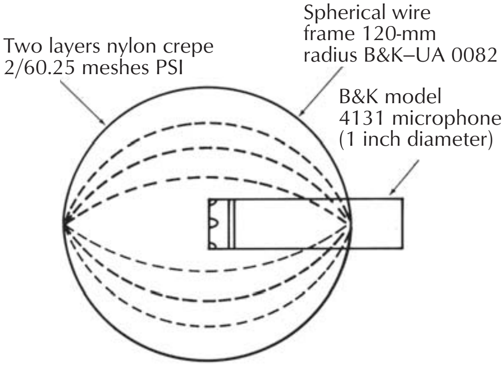


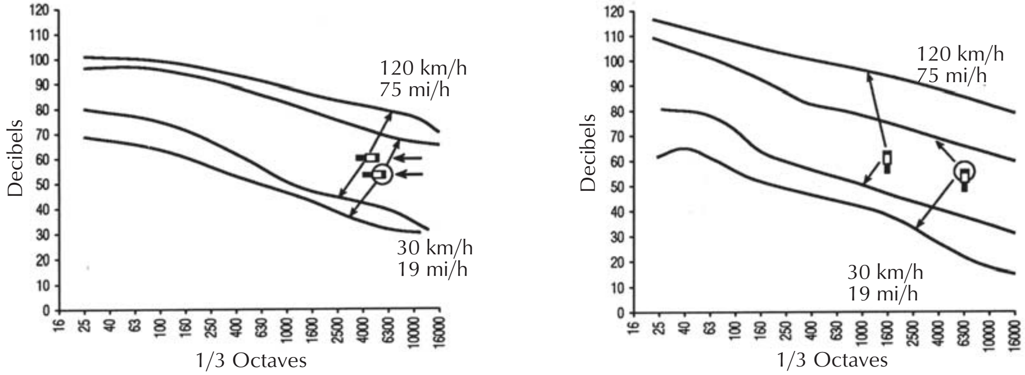
Figure 16-163. Typical silk-covered windscreen and microphone. Courtesy B and K Technical Review.

16.12.2.1 Wind Noise Reduction Figures for Rycote Windshielding Devices

Rycote has developed its own technique for measuring wind noise that uses real wind and a real time differential comparison. The technique compares the behavior of two microphones under identical conditions, one with a particular wind noise reduction device fitted and the other without, and produces a statistical curve of the result corrected for response and gain variations.

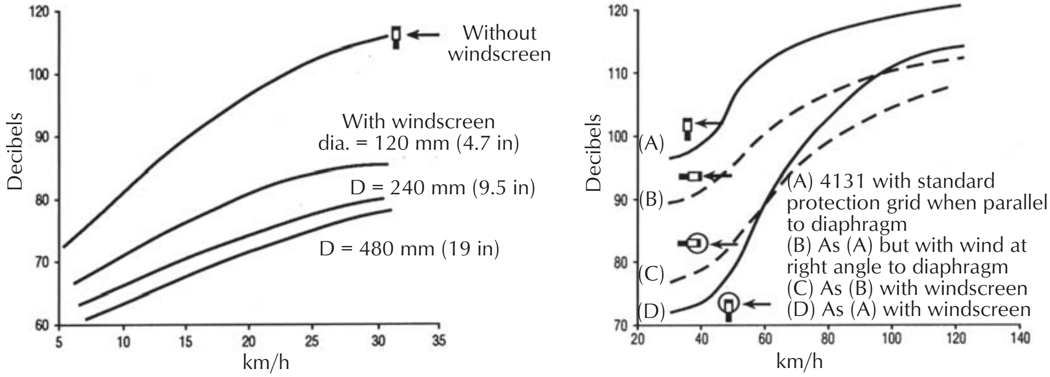
Fig. 16-165 is a Sennheiser MKH60 microphone—a representative short rifle microphone—without any low-frequency attenuation in a wideband (20 Hz–20 kHz) test rig.

When a wind noise reduction device is fitted, its effect on the audio response is a constant factor—if it causes some loss of high frequency, it will do it at all times. However, the amount it reduces wind noise depends on how hard the wind is blowing. If there is a flat calm it will have no beneficial effect and the result will be a degradation of the audio performance of the



A. Wind noise as a function of frequency with the wind direction parallel to the capsule.

B. Wind noise as a function of frequency with the wind direction at right angles to the membrane.



C. Wind noise measured with different sizes of windscreens.

D. Wind noise as a function of wind speed in the range 20 Hz to 20 kHz.

Figure 16-164. The effectiveness of the windscreen shown in Fig. 16-163. Courtesy B and K Technical Review.

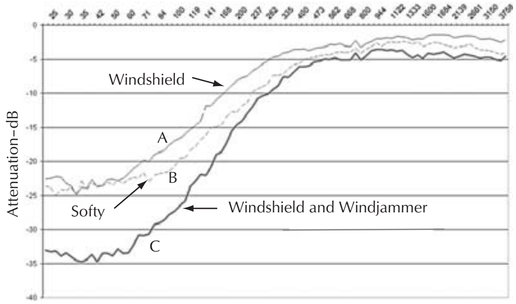


Figure 16-165. Wind noise reduction options for a Sennheiser MKH60 microphone under real wind conditions. Courtesy Rycote Microphone Windshields LTD

microphone. However, in a strong gale a small deviation from a perfect flat response may be insignificant for a >30 dB reduction in wind noise. Wind noise is in the low-frequency spectrum. For a naked Sennheiser MKH60, the wind-produced energy is almost entirely below 800 Hz, rising to a peak of 40 dB at about 45 Hz. It is the effect of a shield at these lower frequencies that is most important. Cavity windshields inevitably produce a slight decrease in low-frequency response in directional microphones but this is not usually noticeable. Basket types have very little effect on high frequency. Fur coverings, while having a major effect in reducing low-frequency noise, will also attenuate some high frequency.

Adding the low-frequency attenuation available on many microphones or mixers (which is usually necessary to prevent infrasonic overload and handling noise when handholding or booming a microphone) may give extra wind noise reduction improvements of >10 dB at the cost of some low-frequency signal loss.

The standard (basket) windshield shows up to 25 dB wind noise attenuation at 35 Hz while giving almost no signal attenuation, Fig. 16-163.

The Softie Windshield is a slip-on open cell foam with an integral fitted fur cover. The Softie reduces wind noise and protects the microphone. It is the standard worldwide in TV. A Softy attenuates the wind noise about 24 dB, Fig. 16-163B.

Adding a Windjammer (furry cover) to the basket windshield will give an improvement of about 10 dB at low frequency to –35 dB, Fig. 16-163C. The attenuation of the Windjammer is approximately 5 dB at frequencies above 6 kHz although this will increase if it is damp or the fur is allowed to get matted. Overall this combination gives the best performance of wideband wind noise reduction against signal attenuation. To determine the correct windscreen for microphones of various manufacturers, go to www.microphone-data.com.

Pop protection is best appreciated when close-talking and explosive breath sounds are particularly bothersome. These explosive breath sounds are commonly produced when saying words involving P and T sounds. The phrase *explosive breath sound* is somewhat of a misnomer since these sounds, without amplification, are normally inaudible to a listener.[15]

The electrical output from the microphone is actually the transient microphone response to this low-velocity, high-pressure, pulse-type wavefront. The P and T sounds are projected in different directions and can be shown by saying the letters P and T while holding your hand about 3 inches (7.6 cm) in front of your mouth. Note that the T sound is felt at a considerable distance below the P sound.

For most microphones, pop output varies with distance between the source and microphone, reaching a peak at about 3 inches (7.6 cm). Also the worst angle of incidence for most microphones is about 45° to the microphone and for a glancing contact just at the edge of the microphone along a path parallel to the longitudinal axis.

sE Dual Pro Pop Filter. An interesting pop filter is shown in Fig. 16-166. The sE Dual Pro Pop pop screen is a two-filter device to suit vocal performances. The device has a strong gooseneck with both a standard fabric membrane and a pro metal pop shield on a hinge mechanism. They can be used separately or both simultaneously depending on the application.

In an emergency pop filters can be as simple as two wire-mesh screens treated with flocking material to create an acoustic resistance.

16.12.2.2 Reflexion Filter

The Reflexion Filter by sE Electonics is used to isolate a microphone from room noises hitting it from unwanted directions, Fig. 16-167.

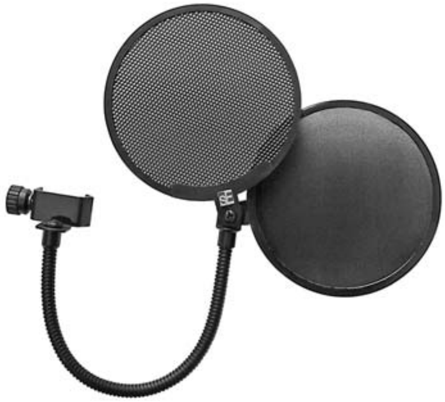


Figure 16-166. sE Dual Pro Pop pop screen. Courtesy sE Electronics.

The reflexion filter has six main layers. The first layer is punched aluminum, which diffuses the sound waves as they pass through it to a layer of absorptive wool. The sound waves next hit a layer of aluminum foil, which helps dissipate energy and break up the lower frequency waveforms. From there they hit an air space kept open by rods passing through the various layers.

Next the waves hit an air space that acts as an acoustic barrier. The sound waves pass to another layer of wool and then through an outer, punched, aluminum wall that further serves to absorb and then diffuse the remaining acoustic energy.

The various layers both absorb and diffuse the sound waves hitting them, so progressively less of the original source acoustic energy passes through each layer, reducing the amount of energy hitting surfaces so less of the original source is reflected back as unwanted room ambience to the microphone. The Reflexion Filter also reduces reflected sound from reaching the back and sides of the microphone. The system only changes the microphone output by a maximum of 1 dB, mostly below 500 Hz.

The stand assembly comprises a mic stand clamp fitting, which attaches to both the Reflexion Filter and any standard fitting shock mount.

16.12.3 Shock Mounts

Shock mounts are used to eliminate noise from being transmitted to the microphone, usually from the floor or table.

Microphones are very much like an accelerometer in detecting vibrations hitting the microphone case. Shock

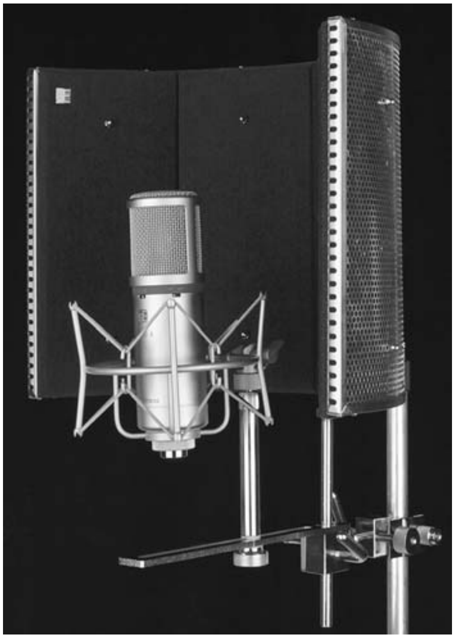


Figure 16-167. Reflexion Filter. Courtesy sE Electronics.

mount suspensions allow a microphone to stay still while the support moves.

Suspensions all use a springy arrangement that allows the microphone to be displaced and then exerts a restoring force to return it to the rest point. It will inevitably overshoot and bounce around, but the system should be damped to minimize this.

As frequency lowers, the displacement wavelength increases so the suspension has to move farther to do the job. For any particular mass of microphone and compliance (wobbliness) of suspension, there is a frequency at which resonance occurs. At this point the suspension amplifies movement rather than suppresses it. The system start to isolate properly at about three times the resonant frequency.

The microphone diaphragm is the most sensitive along the *Z*-axis to disturbances. Therefore the ideal suspensions are most compliant along the *Z*-axis, but should give firmer control on the horizontal (*X*) and vertical (*Y*) axes to stop the mic slopping around, Fig. 16-170.

Suspension Compliance. Diaphragm and so-called donut suspensions can work well, but tend to have acoustically solid structures that affect the microphone's polar response. Silicone rubber bands, shock-cord cat's cradles, and metal springs are thinner and more acoustically transparent but struggle to maintain a low tension, which creates a low resonant frequency, while at the same time providing good *X–Y* control and reliable damping. The restraining force also rises very steeply with displacement, which limits low-frequency performance.

Shock mounts may be the type shown in Fig. 16-168. This microphone shock mount, a Shure A53M, mounts on a standard $\frac{5}{8}$ in – 27 thread and reduces mechanical and vibration noises by more than 20 dB. Because of its design, this shock mount can be used on a floor or table stand, hung from a boom, or used as a low-profile stand to place the microphone cartridge close to a surface such as a floor. The shock mount in Fig. 16-169 is designed to be used with the Shure SM89 shotgun microphone.

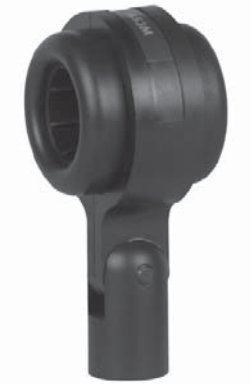


Figure 16-168. Shure A53M shock mount. Courtesy Shure Incorporated.

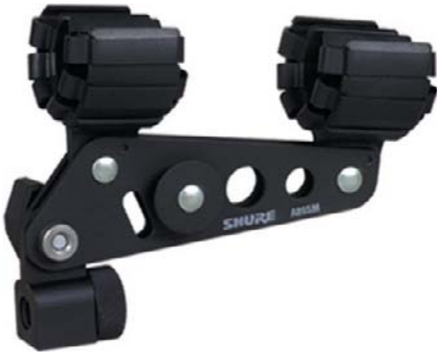


Figure 16-169. Shure A89SM shock mount for a shotgun microphone. Courtesy Shure Incorporated.

Shock mounts are designed to resonate at a frequency at least 2½ times lower than the lowest frequency of the microphone.[21] The goal is simple but there are practical limitations. The resonant frequency (f_n) of a mechanical system can be computed from

$$f_n = \frac{1}{2\pi}\sqrt{\frac{Kg}{w}} \tag{16-32}$$

where,

K is the spring rate of the isolator,

g is the acceleration due to gravity,

w is the load.

A microphone shock-mount load is almost completely determined by the weight of the microphone. To obtain a low-resonant frequency, the spring rate or stiffness must be as low as possible; however, it must be able to support the microphone without too much sag and be effective in any position the microphone may be used.

The Rycote lyre webs rely primarily on their shape to give different performance on each axis. Typically, a 100 g force will barely move a microphone 1 mm along the (up and down) Y-axis, whereas it will move about four times that on the (sideways) X-axis. In the critical Z-axis, it will move almost ten times as far, Fig. 16-170.

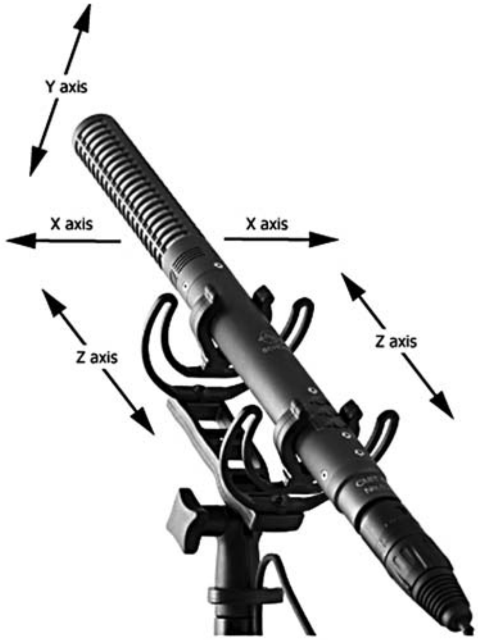


Figure 16-170. A Rycote lyre-type microphone suspension (shock mount) system. Courtesy Rycote Microphone Windshields LTD.

With a very low inherent tension the resonant frequency can be very low, and the Z displacement can be vast. Even with small-mass compact microphones, a resonance of <8 Hz is possible, which means that microphones can be well isolated across almost their entire frequency range.

Damping has to be added to metal spring suspensions, and although integral to rubber band versions, is not very easy to control. With the lyre webs damping can be selected almost independently by choosing a suitable plastic. The Hytrel that Rycote uses not only damps smoothly but maintains its characteristics even down to arctic temperatures. It also has a shape memory that allows it to be tied in eye-watering knots without developing a permanent set—or snapping!

Most suspension systems are difficult to scale.Springs and elastic bands become thin and fragile, and the range of softness for rubber and foam is limited. However, this does not apply to lyre webs. The tiny InVision suspensions, which are visually unobtrusive, isolate compact and similar sized microphones down to <30 Hz, yet are tough enough to be dropped on the floor without risk. Fig. 16-171 shows the actual measured performance of the transfer function for a Schoeps CCM4 microphone being shaken with pink noise in an InVision mount. Trace A shows the output from the microphone with the shaker operating but not touching the mic, revealing the inherent coupling through air and the building itself. Trace B is with the shaker directly coupled to the microphone body to reveal the actual level of vibration input. Finally, the trace C shows the microphone's output with the shaker knocking the bar of the mount, thus demonstrating the effectiveness of the suspension.

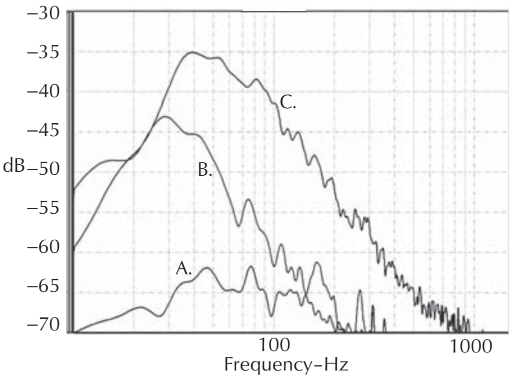


Figure 16-171. Effectiveness of an InVision mount. Courtesy Rycote Microphone Windshields LTD.

To determine the correct suspension systems for microphones of various manufacturers, go to www.microphone-data.com.

16.12.4 Stands and Booms

Microphones are mounted on microphone floor stands or table stands to place the microphone in front of the sound source. The floor stands are usually adjustable between 32 and 65 inches (0.8–1.6 m) and incorporate a ⅝ inch – 27 thread for mounting the microphone holder or shock mount. They normally have a heavy base or three widespread legs for stability.

The table stands are 6–8 inches (15–20 cm) high and often incorporate a shock mount and an on–off switch, as shown in Fig. 16-172.

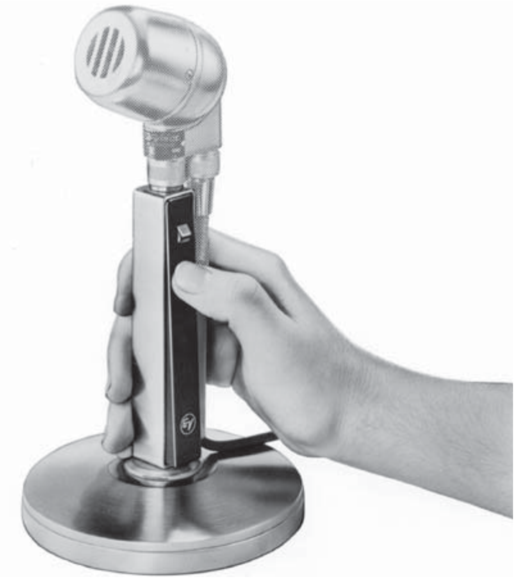


Figure 16-172. Electro-Voice table microphone stand with push-to-talk switch. Courtesy Electro-Voice, Inc.

Small booms, which are mounted on the standard microphone floor stand, are normally used to put the microphone in a place where it is difficult to reach with a floor stand, Fig. 16-173. They are also useful when micing from above the source. Combination booms and stands are often on wheels or flat tripod legs and adjustable from 60–90 inches (1.5–2.3 m) vertically and 90–110 inches (2.3–2.8 m) horizontally, Fig. 16-174.

It is important that the boom and/or microphone stand be easily adjusted and that the clutch/brake system has a positive lock. Better microphone stands incorporate a piston-type air suspension system for effortless height adjustment and microphone protection.

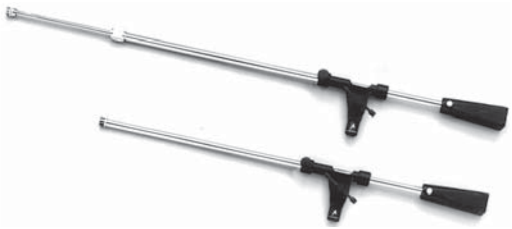


Figure 16-173. Atlas BB-44 microphone boom. Courtesy Atlas Sound.

Large booms, as used in television and motion-picture sound stages, are motorized and often include a stage for the microphone sound person.

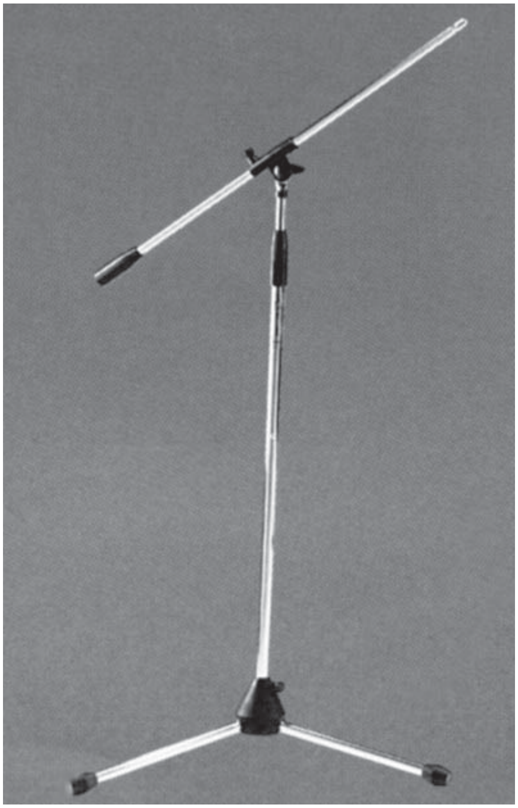


Figure 16-174. Adjustable microphone stand/boom. Courtesy Atlas Sound.

16.12.5 Attenuators and Equalizers

Attenuators, equalizers, and special devices from Electro-Voice, Shure, and others are available to reduce the microphone output level or shape the response to roll off the low or high end, increase the 3–5 kHz articulation region, or reverse polarity. These units normally have standard input and output male and female XLR or ¼ inch phone plug connectors. Attenuators are also available to be installed between the capacitor capsule and the condenser microphone electronics to eliminate overload from high-level sources.

16.13 Microphone Techniques

Micing is more of an art than a science. Therefore there is no one way to position a microphone for good recording. It is subjective and at the control of the engineer. The discussions of microphone placement in the following sections are only suggestions or the ideas of one engineer.

The quality of the reproduction can be greatly influenced by the position of a microphone in relation to the sound source. When only one microphone and one sound source are involved, this positioning is fairly straightforward: the closer the microphone, the more the direct sound will dominate over the reverberant sound. Except in an anechoic chamber, there will always be a certain amount of reflected sound present in the microphone output. This results from sound bouncing off boundaries such as the floor, ceiling, walls, and objects of significant proportions located in the area of the microphone. At a certain distance from the sound source, the amount of reflected sound will exceed the amount of the direct sound. The microphone is then said to be in the reverberant, or far, field. The effect is to make the acoustic environment (usually a room) more evident to the listener than would be the case with close micing (microphone in the near field).

The proper position of the microphone depends on the effect desired. Close micing produces a highly present, up-front sound, with little of the acoustic environment evident, whereas distant micing produces a more spacious sound with the room characteristics becoming very obvious. A close microphone position may not accurately reproduce the sound of the source, and equalization may be required to achieve a sound similar to the natural sound. If the room acoustics are not suited to the sound reproduction desired, a distant microphone position may produce an unpleasant or unintelligible result. The correct choice requires the engineer to choose the appropriate microphone position for the sound desired. A microphone placed an inch from a snare drum will produce an up-front, bigger-than-life sound, which could be appropriate for a modern rock recording but might be totally inappropriate for a jazz or big band recording. Distant micing of the snare drum could produce a powerful effect, in any kind of music, since the contribution of a good room might be important to the music.

It is rare that there is just one microphone and one sound source. Modern recording often requires the use of multiple microphones. Microphone placement then becomes more complicated, because as the microphone is moved farther from its intended source, more of the other sources will be picked up as well. No instrument is a point source, and there are different characteristic sounds emanating from various places on the instrument (i.e., a flute has vastly different sounds coming from the open end, the body of the flute, or the mouthpiece). Most instruments have complex directional characteristics, that vary from note to note. Even instruments of the same make and model can sound quite different from one another.

Whenever there is more than one microphone receiving sound from a single source, a problem of time and phase differences can become audible. This problem can have a major effect on the frequency response, presence, and clarity of the recording. The result for spaced microphones can be a comb filter effect, which will tend to reduce presence, upset the natural balance of various notes and overtones, and disturb localization of the source. In an extreme case, certain notes may be attenuated to inaudibility. In practice, the contribution of room reflections, pickup by other microphones, and intrinsic instrument imbalances may mask many of these effects.

Multitrack recording generally requires the engineer to isolate instruments so that only the intended source is recorded on each track. Sometimes this is simple because the track is being overdubbed and only that one instrument is in the studio. At the other extreme, an entire ensemble may be playing at once, yet the situation may require that all instruments be totally isolated on the tape tracks so that they can be individually mixed, processed, or even replaced with no effect on the other instruments. The latter requires very careful microphone choice and placement and/or the use of isolation booths for some troublesome instruments. If the musical balance is good in the room, the job is fairly simple. But if there are obviously incompatible instruments playing simultaneously (e.g., heavy drums versus a finger-picked acoustic guitar), isolation solely through microphone technique becomes next to impossible.

16.13.1 Stereo Micing Techniques

Modern recording practice often employs multiple microphones, each feeding a separate track of a multitrack tape machine. Sound reinforcement practice usually requires good isolation of the various sound sources. In either case, the end result is a composite of a number of monaural sources, which are often placed in the stereo image with pan pots. This practice is not the same as true stereo recording, which can provide a sense of depth and realism unachievable with panned mono sources. It requires greater effort for superior results; a good acoustic environment is essential.

There are a number of stereophonic recording techniques available to the engineer. The simplest requires two microphones, often omnidirectional types, spaced apart by a distance ranging from several feet to more than 30 ft (9 m), Fig. 16-175. The spacing depends on the size of the sound source, the size of the room, and the effect desired. A broad source like an orchestra will require a wider spacing than a small source such as a single voice or instrument. If the microphones are too far apart, a hole in the middle of the stereo image will result, since the sound produced in the center of the stage will be too far from either microphone. When placed too closely together, a mono result will be obtained. When the spacing is comparable to the wavelength of the sound, phase cancellations may result (comb filters), which will destroy the monaural compatibility of the recording. The best spacing seems to be from 10–40 ft (3–12 m). Experimentation is necessary since every situation will be different. Needless to say, good monitoring is required; stereo headphones will not generally reveal defects evident on good monitor loudspeakers. A method of summing the two channels to mono is essential for testing compatibility.

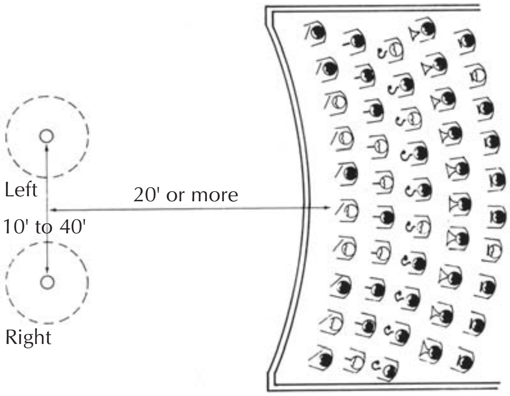


Figure 16-175. Spaced omnidirectional microphones for stereo recording.

Variations on the spaced microphone technique involve using bidirectional or unidirectional microphones, which may be helpful when the room characteristics are not perfect for the material being performed. Adding a microphone in the center, fed to both left and right channels (*fill* microphones), and combinations of spaced micing and other techniques might be required.

16.13.2 Microphone Choice

Every microphone type has certain characteristics. These characteristics must be taken into account when choosing a microphone for a specific application. Some of the factors to be considered are general type (condenser, moving coil dynamic, ribbon); directional pattern (omni-, bi-, or unidirectional); and specific microphone traits (bright, bassy, dull, presence peak, and so on).

Also, the susceptibility of the microphone to overload or its tendency to overload the associated preamplifier must be considered. The off-axis frequency response can have a large effect on the sound of a microphone in a particular application. Certain microphones may exhibit unusual traits that may make them more, or less, suitable for a certain application. For example, the design of the grille may have a major effect on the sound of a microphone when recording closely micing vocals.

Some of these characteristics can be inferred from the microphone specifications (i.e., frequency response, overload point, directional pattern—both on-and off-axis). Other characteristics are not as easy to measure or visualize, and experience and experimentation are necessary to make an intelligent choice.

16.13.3 Microphone Characteristics

There are many criteria used to judge the suitability of a microphone for a particular application—some are quite subjective. Frequency response is one obvious characteristic, distortion is another. The ability of a microphone to accurately translate waveforms into electrical signals is vital for good reproduction. Generally, the less massive the internal parts that must be moved by the sound pressure, the more accurate the reproduction, especially the reproduction of waveforms with steep leading edges and/or rapid level changes (e.g., percussive sounds). The condenser microphone has the lowest mass (only a thin plastic diaphragm with a very thin coating of metal must be moved by the sound pressure). The diaphragm and coil in the dynamic microphone have considerably more mass than the condenser diaphragm. The ribbon in a ribbon microphone has relatively low mass and is somewhere between the condenser and the dynamic microphone.

It would seem that the condenser microphone would always be the best choice, but other factors must be considered. Condenser microphones are generally less rugged than dynamic ones, and since they are usually more expensive, the decision to place a valuable microphone in a position where it could be hit or knocked over must be weighed against the possible benefit of improved sound. Also, condenser microphones contain internal active electronics, which can be overloaded by high sound levels. Many condenser microphones contain switchable or insertional pads, but long before the overload distortion becomes apparent, clipping of the transient peaks may muddy the sound in a subtle way.

Ribbon microphones are somewhat fragile. They can be especially vulnerable to blasts of air that can occur when closely micing vocals, inside a bass drum, or even when a door is slammed in an airtight studio.

In each type of microphone, there are many other factors that can affect the sound. The design of the mounting for the microphone components, the internal obstacles in the sound path, and the effect of the body of the microphone, all can have a major effect on the ultimate sound reproduction.

16.13.3.1 Directional Pattern

It might at first seem that the unidirectional microphone would be the universal choice for all applications, since picking up the intended source is the goal. It is true that unidirectional microphones (see Section 16.2.3) have the greatest application, but there are situations that require the use of omnidirectional microphones, which are designed to pick up sound from all directions as nearly equally as possible (see Section 16.2.1), or bidirectional microphones, which are sensitive to the front and back, but insensitive to the sides (see Section 16.2.2). But it is possible, in some situations, to obtain greater rejection of unwanted sound with an omni- or bidirectional microphone than would be possible with a unidirectional pattern.

Unidirectional and bidirectional microphones often exhibit a proximity effect, in which the response to lower frequencies (generally below 150 Hz) is emphasized when the microphone is placed close to the sound source (Section 16.2.3). Close may be a couple of inches or a couple of feet, depending on the microphone. Various designs have been developed to minimize or eliminate this effect. A switchable high-pass filter may be included on the microphone to roll off the bass in close micing positions. Proximity effect must be considered when choosing and placing a microphone. Sometimes the effect can be used to advantage (i.e., when additional bass response is desirable, perhaps on a snare drum or on certain vocals). But often the proximity effect emphasizes the (unrelated) tendency of some sound sources to sound more bassy when close mic'ed.

Directional microphones do not have the same frequency response off-axis as they do on-axis. This can cause increased apparent sound leakage from other sources, tonal aberrations of the reproduced sound, or unexpected phase cancellations. For example, many directional microphones exhibit less directionality at both higher frequencies and lower frequencies. If such a microphone were used to close mic a snare drum, the amount of pickup of the nearby bass drum and cymbals might be excessive.

16.13.4 Specific Micing Techniques

There are probably as many methods of using microphones as there are engineers. Contrary to popular opinion, there does not seem to be any special microphone or magical technique for recording any particular sound. What is right is what sounds best. The following discussion is merely a review of some common techniques widely employed and likely to work well in many circumstances.

16.13.4.1 Musicians

The first requirement for obtaining a good sound from any instrument is a superior player. An experienced studio musician can make almost any studio or engineer sound good. Unfortunately, the engineer usually has very little to say about the musicians who are hired for the session. When inexperienced players record, they may often expect to be made to sound like whoever their idols may be. They probably don't want to know that their idol spent the last 10 years or more learning how to use the studio, and they may be likely to blame the engineer for their inability to play properly for recording. There isn't much that can be done in such a circumstance.

16.13.4.2 Drums

Studios involved in music recording are more often judged by their drum sound than by anything else. It is true that much contemporary music relies heavily on drums and that getting the best possible sound is a goal worth pursuing. There are any number of ways to record drums, but the most commonly used technique utilizes close micing.

Just as the musician is a vital element in obtaining a good sound, the drums themselves must be in good condition and properly tuned to obtain their best sound. The type of drum head used will have a major effect on the sound.

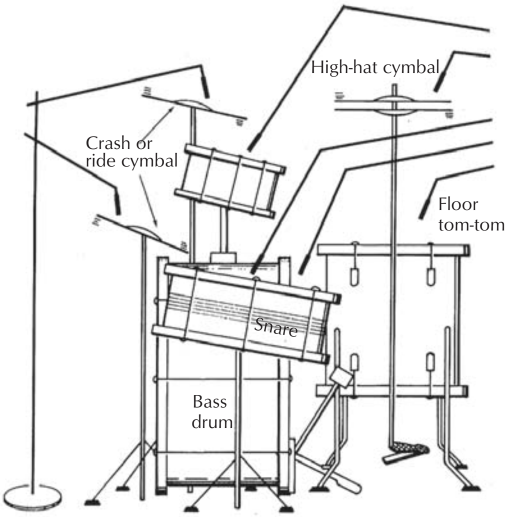
Micing Each Drum. A micing arrangement that is almost standardized requires the use of one microphone on each drum, Fig. 16-176. In addition, one or more microphones may be suspended over the drum set to pick up either an overall sound or primarily cymbals. How closely each microphone is placed depends on several factors: how tight a sound is required, which in turn is related to the relative liveness and character of the room; what isolation problems might exist, in terms of various drums leaking into other drum microphones and leakage from other instruments in the room; how dangerous it may be to place an expensive and fragile microphone in a position of possible destruction by an overly enthusiastic or inaccurate drummer; and whether the microphone and/or console can take the level produced without distortion.

Above versus Underneath Micing. Individual drums can be miced either from above or below, Fig. 16-177. The two positions will usually have vastly different sounds. If the sound is appropriate, the underneath position may be preferable if isolation is a problem.

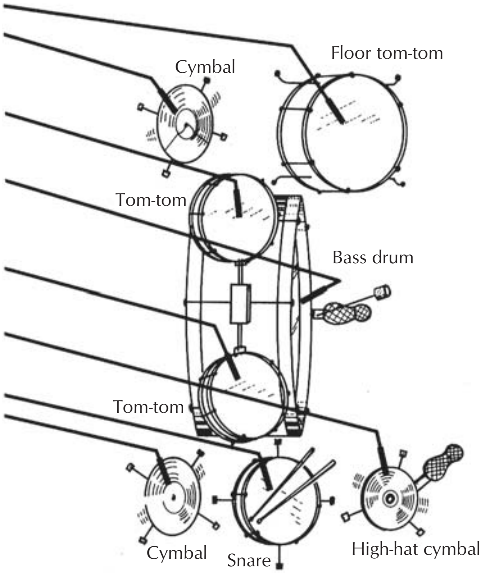
When miced from above, microphones are commonly positioned at an angle to the drum head and near the edge of the drum. Seemingly minor changes in position can have a major effect on the sound, especially with some microphones.

Bass Drum. For recording, bass drums usually have only the beaten head, which is not to say that bass drums with both heads cannot be recorded, however. For some music, the use of both heads is preferable. In the single-head configuration, the usual microphone placement is within the shell of the drum, with the microphone aimed toward the beater, Fig. 16-178. Experimentation is required, however. Closer or farther distances, off-axis microphone positions, or even placement on the opposite side of the head may result in the desired sound.

Tom-Tom Micing. Tom-toms, too, often use only the top head. This facilitates underneath micing. In micing any drum, it is probable that simultaneous top and bottom micing will result in difficulty due to phase discrepancies. The use of phase-reversal switches at the board and minor position adjustment may be required.



A. Side view.

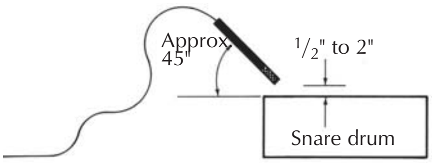


B. Top view.

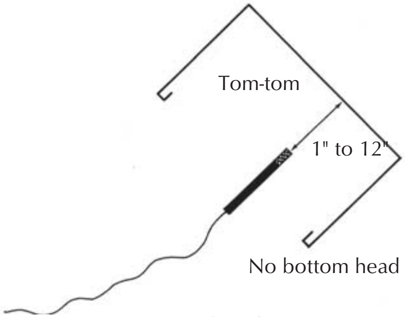
Figure 16-176. Close drum micing (detail).

Cymbals. The high hat and cymbals can be mic'ed from above or below, but the above position is more commonly used. Overhead microphones are often positioned above the entire drum set, usually as a stereo pair. How high they are set will depend on the effect desired; a relatively high placement will provide more of an overall drum sound, with more room characteristics than a closer position.

It is not unusual to pick up sufficient cymbals, or even excessive cymbals, from just the other drum


Approx. 45°
1/2" to 2"
Snare drum

A. Above drum.


Tom-tom
1" to 12"
No bottom head

B. Below drum.

Figure 16-177. Bass drum micing.

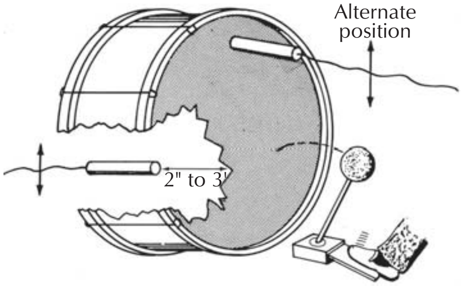
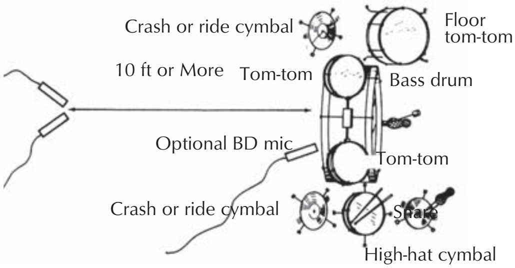

Alternate position
2" to 3'

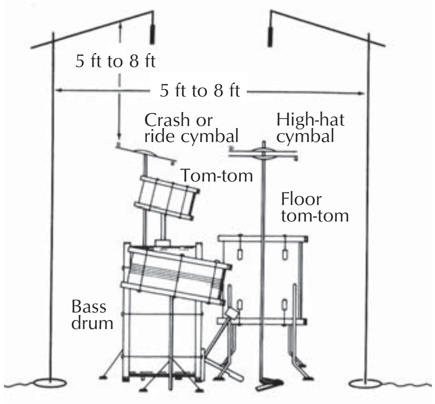
Figure 16-178. Bass drum micing.

microphones without the overhead microphones even being on. The amount of cymbal leakage will be determined mostly by the drummer's technique and balance, with the room characteristics also being a factor.

Other Drum Micing Techniques. Close micing every drum is only one method. Another is to use relatively distant microphones to pick up an overall drum sound, Fig. 16-179. This, of course, results in much more room sound and possible leakage from other instruments. It also requires that the drummer play all the drums, and particularly the cymbals, in the proper balance. The engineer has much less control of the sound. This approach will not be successful in poor rooms, nor with drummers who do not correctly balance their various drums and cymbals. But in a good room, with a good drummer, the sound can be quite natural and often very powerful. A common technique is to use two overhead microphones, placed in such a way as to capture the natural sound and balance of the drum set. Some experimentation will be required to find the proper placement. Usually a separate bass drum microphone is used as well, to give the bass drum better definition and more punch.


Crash or ride cymbal
Floor tom-tom
10 ft or More
Tom-tom
Bass drum
Optional BD mic
Tom-tom
Crash or ride cymbal
Snare
High-hat cymbal

A. Coincident microphones.


5 ft to 8 ft
5 ft to 8 ft
Crash or ride cymbal
High-hat cymbal
Tom-tom
Floor tom-tom
Bass drum

B. Overhead spaced microphones.

Figure 16-179. Distant drum mic'ing.

16.13.4.3 Piano

Pianos are often recorded in stereo and can add width and a greater sense of space if done properly. Multiple microphones spread all over the sounding board may seem like an ideal way to pick up the full piano sound, but this procedure can lead to a very artificial and distant sound when heard in mono.

First, be certain that a stereo piano is really desirable. In multitrack recording, are there sufficient tracks available? And is the piano sound required to be so big? A mono piano can often have more punch and might be a better choice.

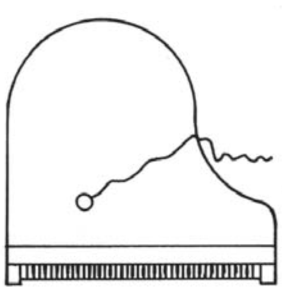
For a mono track, one microphone is usually all that is needed. For stereo, a pair of adjacent directional microphones will probably suffice. In micing either a grand or upright piano, keep in mind that the sounding board and not the hammers and strings is the source of most of the sound. With a good piano, there may be surprisingly little difference in the sound picked up from various areas of the sounding board. A commonly mic'ed point is where the bass and treble strings cross, Fig. 16-180. Variations such as micing from beneath (or in the case of an upright, in back of) the sounding board, inserting microphones into the circular holes in the harp, or using various types of piano pickups can all be tried. Each piano is different, and each player will also have a large effect on the sound, so a variety of techniques should be tried.

The PZMicrophones can be used in recording piano. They can be placed on the inside of the piano lid and the lid closed for improved isolation.

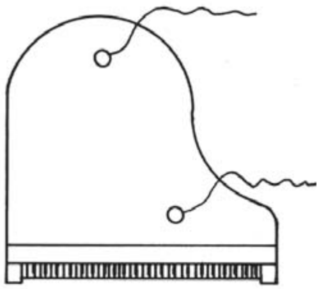
As in all percussive instruments, the peak level produced by a piano can be far greater than the level shown on the volume unit (VU) meter. Peaks 20 dB above the meter reading are common. Since just about everybody knows what a piano sounds like, and since the instrument is so frequently featured in musical pieces, any distortion will be very obvious to the listener. Even a distortion that only occurs on the peaks can be evident as a dulling of the piano attack, a kind of audio blurriness. The peaks can really strain the dynamic range of microphones, preamps, and storage medium. If condenser microphones are used, be sure the pads are switched on even if the level seems moderate. Also, some engineers routinely record piano at a somewhat lower than normal level to avoid tape saturation.

Obtaining satisfactory isolation while still getting a good sound can be a problem with the piano. Isolation can be achieved with a booth, of course, but careful micing and some baffling can often work almost as well. One technique used in many studios is to place the microphone in the piano and then close the lid as much as possible. Often the short-stick position of the lid works well. Then carpeting or other dense, heavy, absorbent material is draped over the piano. With a good arrangement of other instruments and reasonably balanced volumes, very little leakage should exist. Another technique requires that microphones be mounted inside the piano, usually suspended from the lid, in such a way that the lid can be completely closed. The PZM type of microphone is particularly well suited for this approach.

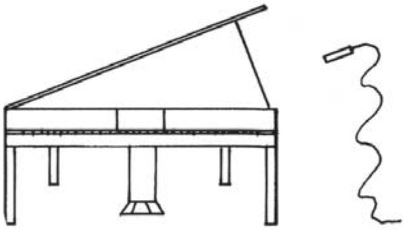
Of course, a much better sound is obtained with the lid open and with perhaps a little more distance between the sounding board and the microphones. Sometimes removing the lid and suspending the microphones above the piano work well. (Most pianos have pins in the



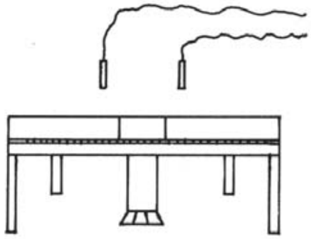
A. Single microphone or coincident pair (typical).



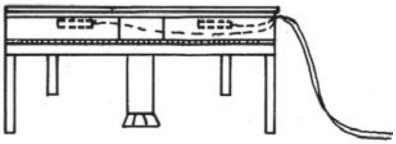
B. Spaced microphone (typical).



C. Distant mic'ing.



D. Overhead mic'ing (piano lid removed).



E. PZM mic'ing – lid closed (placement similar to B).

Figure 16-180. Piano mic'ing.

hinges that can be easily removed for this purpose.) Fairly distant micing may also sound good.

16.13.4.4 Vocals

A single vocal, either speaking or singing, is usually recorded with one microphone placed within 2 ft (0.6 m) of the mouth. For popular music, it is common to have the singer very close to the microphone; in a recording of a classical singing voice, a greater distance is appropriate, even up to several feet may be used if the room acoustics permit. Speakers at a lectern usually are 1–2 ft (0.3–0.6 m) from the microphone.

16.13.4.5 Singers

Although vocals could be recorded in stereo, with any of the techniques previously described, it is customary to record the voice in mono. It is basically a point source, with little directional information. In a superior acoustic environment, such as a good concert hall, natural reverberation may be mixed in with additional microphones. But most often artificial reverb is added. It can be stereo and add considerable depth and width to the voice.

Condenser microphones, placed very close to the mouth, are the usual choice in the studio. A pop filter will be necessary for all but the most careful singers. This prevents explosive sounds from being produced when the vocalist sings a word containing Ps or other hard consonant sounds. It is important to remember that the output level of the microphone will adhere to the inverse square law, which states that if the distance from the source to the microphone is doubled, the level will be reduced to one-quarter. Experienced vocalists are well aware of this phenomenon and may even use it to obtain certain effects. The inexperienced or inattentive singer will probably require electronic processing (i.e., limiting) to obtain a satisfactory performance. This problem is further complicated by the trend toward mixing vocals quite low in the musical track and relying on processing to maintain intelligibility.

In the studio, it is often necessary to provide an acoustic environment less reverberant than normal for the recording of vocals. Cutting down on reverberation could be accomplished with a separate vocal booth with highly sound-absorbent surfaces, or it can be obtained by placing absorbent baffles around the singer and microphone in the studio, Fig. 16-181. On the other hand, it may sometimes be necessary to emphasize the reverberation for a special effect by distant micing or by mixing in another microphone placed some distance away.

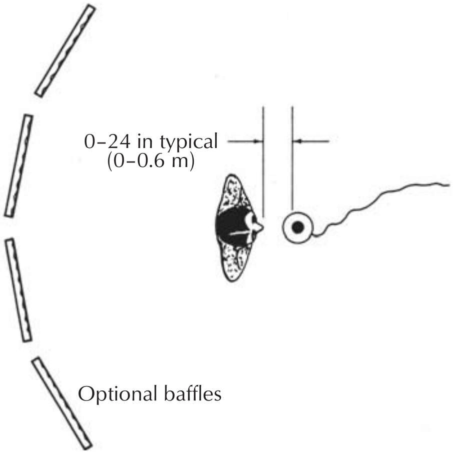


Figure 16-181. Vocal mic'ing.

Proximity effect can be a problem with vocals. Many microphones have provision for a bass roll-off, which can be used to correct this deficiency. This approach is often superior to using equalization in the control room, especially if a limiter is used before the equalizer (the limiter would respond to the emphasized bass and thus not accurately track the vocal intensity). Some singers prefer the effect obtained from proximity, using the bass boost in their performance to emphasize certain words or phrases.

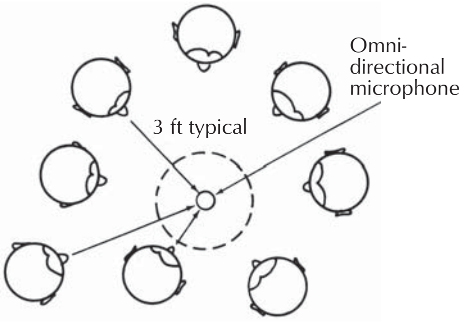
In a live performance, large studio condenser microphones would be inappropriate. With their suspensions and pop filters and the large microphone stand required, they would obscure the singer's face. What is needed is a relatively small, rugged microphone that can be handheld if desired. Although there are a number of condenser microphones that can be used this way, the usual choice is a compact dynamic microphone with built-in pop filters, integral shock mounting, and switchable bass roll-off.

Good directionality is required of a live performance microphone. The usual practice of providing the singer with a stage monitor loudspeaker, usually placed within a few feet of the microphone, requires good rejection of sound from off axis to minimize the possibility of feedback and reduce the degradation of the sound from the vocal microphone picking up the monitor's reproduction of the other instruments and voices. Some microphones designed for live work have their direction of minimum sensitivity oriented toward the direction where the most unwanted sound would come from (i.e., not directly off the back of the microphone, but at some intermediate angle).

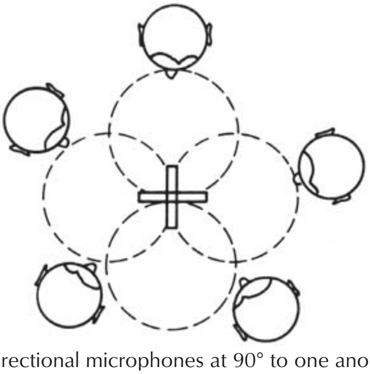
16.13.4.6 Group Vocals

A vocal group could consist of two singers or a chorus of several hundred. For a small group (less than eight), a single microphone with an omnidirectional pattern placed in the center of a circle of vocalists often works well, Fig. 16-182. This microphone arrangement requires that the singers achieve a proper balance of voices in the studio. The final balance can be fine tuned by having the necessary voices move closer to or farther from the microphone. If the singers are relatively close to the microphone (two feet or less), then their positions become more critical. A small change in position can have a major effect on the blend.

For stereo, the group could be divided into two circles, each with its own omnidirectional microphone in the center. Two bidirectional microphones, oriented at 90° to one another and placed one above the other, could be used to obtain a stereo omnidirectional recording when placed in the circle of vocalists, Fig. 16-183.



Adjust spacing to obtain proper balance of voices
Figure 16-182. Group vocals micing–monaural.



Bidirectional microphones at 90° to one another
Figure 16-183. Group vocals micing stereo.

Whenever omnidirectional microphones are used, the room becomes more apparent in the recording than it would with unidirectional microphones. This effect must be considered when recording group vocals in this manner.

If greater presence is required (or less room sound) or if the balance must be controlled by the engineer for some reason, individual microphones could be used for each singer; however, this method has obvious practical limitations if the group is large. It also requires more set-up and balancing time, puts a musical burden on the recording personnel, and might have a disappointing result if lack of isolation creates phase problems when mixing the multiple microphones.

For really large groups, techniques similar to those described for string sections might be employed.

Typically, group vocals will be recorded as an overdub on a previously recorded musical track, requiring the vocalists to wear headphones. With a number of singers wearing headphones (which could be turned up quite loud) standing next to an omnidirectional microphone, a significant amount of leakage from the headphone mix is possible. This leakage from the headphone mix can become even more of a problem if one or more of the singers prefer to remove one side of the headphones from his or her ear in order to better hear their own voice and/or the blend of the other voices. Background vocals are often by nature relatively quiet parts requiring higher than normal gain on the microphone channel. All these factors can combine to degrade the entire recording seriously.

Solutions might be to use as low a headphone level as possible, have the singers sing as loudly as is appropriate for the part, turn the microphone off when the vocalists are not singing, or use a noise gate to do this automatically. In a really severe situation, the solution might be to use individual directional microphones.

16.13.4.7 Lectern Microphones

For redundancy, two or more microphones are often provided on a lectern. Only one should be active at a time, or phase cancellations can result. Often two microphones are arranged on opposite sides of the lectern, angled in toward the talker. The goal is satisfactory pickup as the speaker moves from side to side. This arrangement can cause serious phase cancellation problems because of the spacing (usually a couple of feet) resulting in feedback problems since the normal frequency response has been disturbed through the comb filter effect. A better arrangement places the two microphones in the coincident configuration as close together as possible and angled toward opposite sides of the lectern, Fig. 16-184. The outputs can be summed with no phase problems. The angle between the microphones may be changed from the normal 90° if necessary to obtain proper coverage.

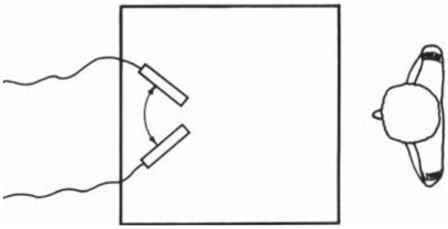


Figure 16-184. Lectern microphones for increased coverage pattern.

16.13.4.8 Strings

Although strings could be close miced, this approach usually results in an unnatural sound. Distant micing is more appropriate but puts a greater demand on the room acoustics. Obtaining a good string sound really requires a good room of considerable size.

A string quartet might sound fine recorded in a relatively small studio (2200 ft^3 or 62 m^3), but a large string section needs more volume. Not only will a larger room accommodate more players, but the microphone placement will also be simpler and the results will be closer to the actual sound of the section.

Each instrument could have a microphone, and this would give the mixer complete control of the balance of all the strings. But unless a great deal of time is available to obtain the proper balance, this approach is not cost effective when recording highly paid musicians. It does not guarantee the best results, either.

At the opposite extreme, a single microphone placed at a point determined to provide the best overall balance and sound could be a simple and quick way to get good results, Fig. 16-185. This placement works well if the engineer is familiar with the room and can rapidly duplicate a setup that has been successful in the past. A coincident pair can provide the same sound if stereo is required.

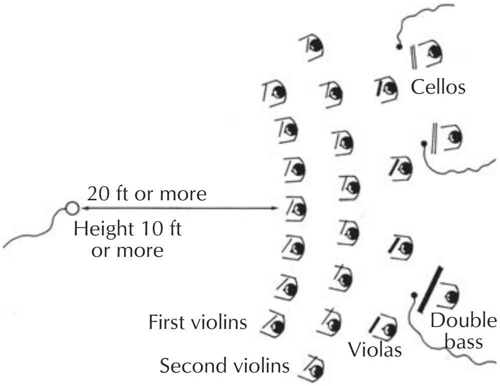


Figure 16-185. Single microphone (or coincident stereo) for string section recording. Optional microphones are shown for cello and double bass.

Another technique is to mic the ensemble in sections, Fig. 16-186, providing, for example, a single microphone for the first violins, another for the second violins, another for the violas, and so on. Cello and double bass often have microphones to pick them up individually in this type of setup.

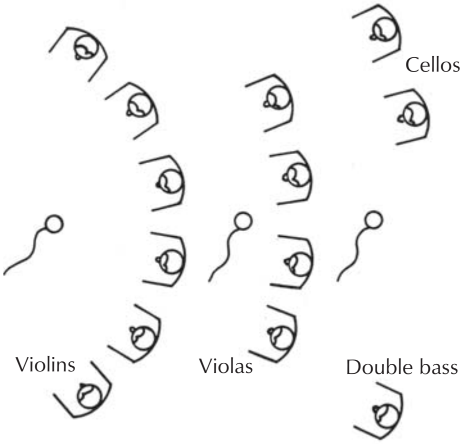


Figure 16-186. String micing by section.

It is also possible to set up microphones above each row of players, or above each two rows. This method is often used in conjunction with the single overall microphone.

In a practice session, Fig. 16-187, the setup is often a composite of all of these techniques: a single coincident pair at a distant point (perhaps 15–20 ft [5–6 m] from the first row, and up as high as practical in the room); a set of microphones over each section (one microphone for every two players, up above the space required for their bows and slightly in front of the instrument); and individual microphones for the cellos and basses (a foot or two in front of the instrument, opposite the F holes). At the start of the session, the overall microphone would first be monitored to determine what, if any, balance problems exist. If time permits, the overall microphone position might be changed to obtain a better balance. If the desired balance cannot be obtained with the single microphone, the necessary individual section microphones may be brought into the mix. In many practically sized rooms, it is not possible to obtain a good balance of the near strings (usually violins) and the far strings (cello and bass). Careful use of the section microphones can correct this.

Since good high-frequency and transient response is required to reproduce the string section sound, condenser microphones are the most frequent types used for string recording.

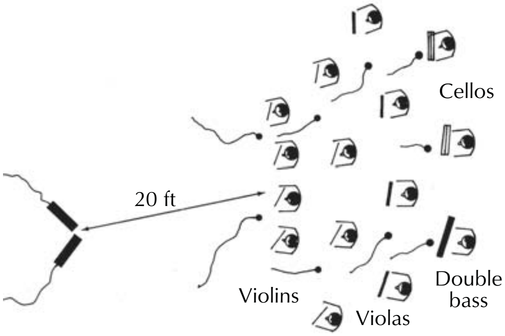


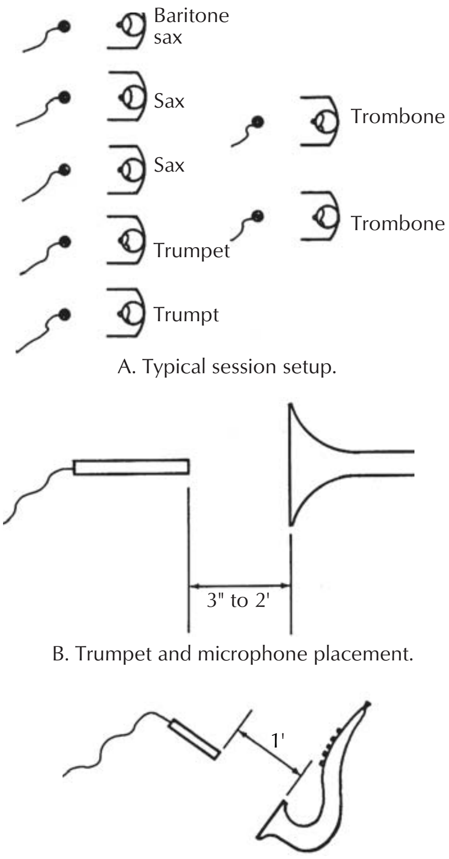
Figure 16-187. Typical composite technique for string micing .

16.13.4.9 Horns

In the recording world, horns are any brass instrument: trumpets, trombones, saxophones, and so on. Modern recording of popular music usually requires close micing of individual instruments, Fig. 16-188. Since many horns are capable of producing very high sound-pressure levels (as high as 130 dB), it is important to choose microphones that will not be overloaded by this close placement. Also, pads may be required to prevent overloading the mixer preamplifier or saturating an input transformer.

Condenser microphones are often used to pick up horns, but ribbon and dynamic types may also give good results.

It is important to remember that the sound produced by these instruments does not come entirely from the bell; this is particularly true of saxophones. Although the instrument output may be loudest at the bell, the contribution of the various other parts of the horn cannot be ignored. The microphone position is often a compromise between the presence of very close placement, the better tonality of a slightly greater microphone distance, the leakage from other instruments as the microphone distance is increased, and the degree of room contribution desired in the finished recording. Depending on the effect desired, 6 inches to a couple of feet may be appropriate.



A. Typical session setup.

B. Trumpet and microphone placement.

C. Saxophone micing.

Figure 16-188. Horn and microphone placement.

16.13.4.10 Woodwinds

Instruments like the oboe, flute, bassoon, clarinet, and their variations cannot generally have microphones placed too closely and still retain their character. In popular music they are often mic'ed individually at a distance of one to several feet, which is generally not far enough to provide a true sound, but the result is often acceptable—or even desirable—for compatibility with other instruments in the song.

Condenser or ribbon microphones are the usual choice. Most woodwinds tend to sound most natural when mic'ed from about 3 ft (1 m) away, with the microphone directed toward the middle of the instrument, or perhaps pointing slightly toward the bell or end of the instrument. Low placement, even on the floor with a PZMicrophone, tends to sound better than high placement.

For classical recording, a more distant pickup is necessary. A woodwind ensemble might be successfully recorded using the techniques described previously for string sections.

16.13.4.11 Electric Instruments

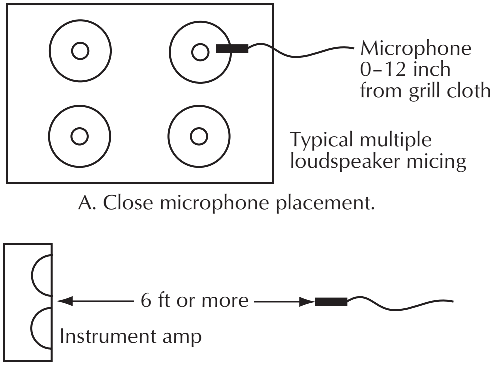
In this category are all instruments designed to be reproduced through amplifiers and loudspeakers. Electric guitar; electric bass; various synthesizer, organ, and other electronic keyboards; and acoustic instruments with attached microphones or pickups designed for amplification all fall into this category.

Generally, these instruments require microphone placement with the associated amplifier/loudspeaker combination. However, another technique is possible and in many cases preferable—that is, the direct recording of the instrument. Since most of these instruments produce a microphone-level high-impedance unbalanced output, all that is required in most cases is a high-quality transformer, providing the match between the instrument and the low-impedance balanced inputs of most mixers. Various direct boxes are available, some with active electronics to provide the required impedance transformation. Almost all provide an output to drive the instrument amplifier as well as the mixer, and most have a ground switch to select the grounding configuration with the least noise.

In many situations, the instrument and its amplifier constitute a system. The amplifier, which may contain loudspeakers or may be connected to a separate loudspeaker system, may have a major effect on the sound of the instrument. Taking a direct feed may result in a totally unnatural sound.

Mic'ing the instrument amplifier may seem simple, but often the cabinet contains several loudspeakers. These may be identical loudspeakers or separate drivers for various frequency ranges. A single close microphone may not provide the proper balance. Even in systems with identical loudspeakers, careless microphone placement may result in phase discrepancies producing a distant and/or colored sound. Two solutions are practicable: either give the microphone a more distant placement, far enough to be equidistant from all the loudspeakers, or position it very close, to pick up only one loudspeaker, Fig. 16-189.

Systems with multiple drivers for different frequency ranges will have to be mic'ed from far enough away that the various drivers are properly balanced. Although it may be possible to mic the individual drivers and mix them for the proper balance, this approach is more prone to error.



A. Close microphone placement.

B. Distant microphone placement.

Figure 16-189. Mic′ing of electric instrument loudspeakers.

Distant micing is often desired, especially for an electric guitar. Naturally, the character of the room must be appropriate.

Often a combination of the direct and mic'ed sound is used. This combination can be effective, but the phase relationship between the two sources will be arbitrary, which can cause severe coloration of the sound. The tonal balance will change unpredictably as the ratio of direct and mic'ed sound changes. This change usually precludes any gain riding of the individual inputs. A phase reversal switch can sometimes be used to optimize the gross phasing between the two inputs.

Instruments like synthesizers or other electronic keyboards generally should be recorded directly. The sound of these instruments is usually not augmented by the addition of a musical instrument amplifier. There are exceptions, however, and the choice of technique depends on the effect desired—perhaps the limited frequency response and soft distortion of a tube-type amplifier is appropriate.

16.13.4.12 Percussion

The most common percussion instrument is the drum kit. Other percussion instruments, such as congas, tympani, handclaps, tambourines, timbales, wood blocks, claves, or maracas, etc., require care in micing due to the extreme levels encountered. It is not uncommon to have levels of +10 dBm and more (open circuit) on the output of a condenser microphone when placed close to a percussion instrument or a piano.

Such levels can be very demanding of microphone electronics, in the case of condenser microphones and the associated mixer. The use of internal microphone pads is essential. Additional padding may be necessary between the microphone output and the mixer input.

The correct micing procedure for percussion instruments depends on the effect desired. A distant micing position is often justified when the sound of the room reverberation adds to the effectiveness of the instrument. Tambourine and handclaps often benefit from the sound of a good room. The resultant sense of space can produce better depth in the recording, and/or the explosive nature of a large, live room can add tremendous punch to the part.

On the other hand, the highly present sound of close micing might be more appropriate in another musical situation. Close, in this sense, might range from fractions of an inch to a couple of feet. Handheld instruments, like claves, must be played at a uniform distance from the microphone, which becomes more critical as the distance decreases.

16.13.5 Conclusion

It is important to remember that there is never only one way to position microphones. The techniques presented here are representative of the methods widely used in the recording and sound-reinforcement industries, but such practices have evolved over many years. Some are traditional; however, there may be better ways. Using the procedures outlined will result in reasonably accurate reproduction, or commercial reproduction as it applies to mainstream music recording. Since sound reproduction can be a creative endeavor, experimentation may yield new techniques. The exact reproduction of the original sound may not be the goal. Perhaps the engineer is attempting to obtain a previously unheard sound or effect. When the luxury of experimentation is available, the engineer may well use the time to pioneer new techniques that can supplement or even replace existing procedures.

Acknowledgments

Thanks to Michael Pettersen, Shure Incorporated, for his assistance in updating and correcting this chapter.

References

1. A.M. Wiggins, “Unidirectional Microphone Utilizing a Variable Distance between the Front and Back of the Diaphragm,” *ASA,* Vol. 26, no. 5, September 1954. (Parts copied with permission.)
2. R. Schulein, Shure Incorporated, “Development Consideration of a Versatile Professional Unidirectional Microphone,” *Journal of the AES,* Vol. 18, no. 1, p. 44, February 1970. (Parts copied with permission.)
3. W. R. Bevan, R. B. Schulein, and C. E. Seeler, Shure Incorporated, “Design of a Studio-Quality Condenser Microphone Using Electret Technology,” *J. Audio Eng. Soc.,* vol. 26, no. 12, p. 947, December 1978.
4. H. Tremaine, *Audio Cyclopedia,* Indianapolis, IN: Howard W. Sams & Co., Inc., 1969, pp. 148-150.
5. Ibid., pp. 150-151, 152.
6. A. P. G. Peterson and E. E. Gross, Jr., *Handbook of Noise Measurements*, General Radio Co., p. 33.
7. Don and Carolyn Davis, *Sound System Engineering,* Indianapolis, IN: Howard W. Sams & Co., Inc., 1975.
8. Henning Gerlach, “Stereo Sound Recording with Shotgun Microphones,” *Sennheiser News*. (Parts copied with permission.)
9. Yuri Shulman, “Reducing Off-Axis Comb-Filter Effects in Highly Directional Microphones,” *J. Audio Eng. Soc.,* Vol. 35, no. 6, June 1987.
10. Y. Ishigaki, M. Yamamoto, K. Totsuka, and N. Miyaji, Victor Company of Japan, Ltd., “Zoom Microphone,” *AES preprint 1718 (A-7)*. (Parts copied with permission.)
11. Manfred Hibbing, “XY and MS Microphone Techniques in Comparison,” *Sennheiser News.* (Parts copied with permission.)
12. Wesley L. Dooley and Ronald D. Streicher, “M-S Stereo: A Powerful Technique for Working in Stereo,” *AES preprint 1792 (J-6)*.
13. David A. Ross, “A Practical Approach to Applying the MS Stereo Microphone,” Shure Brothers, Inc.
14. C.P. Repka, “A Guide to Coincident Mikes,” *Audio Magazine,* November 1978.
15. Klaus Genult and Hans W. Gierlich, “Investigation of the Correlation between Objective Noise Measurement and Subjective Classification,” HEAD Acoustics GmbH, Aachen, Federal Republic of Germany.
16. Dr. Klaus Genuit and Wade R. Bray, “Binaural Recording for Headphones and Speakers,” *Audio Magazine,* December 1989.

17. H.W. Gierlich and K. Genuit, “Processing Artificial-head Recordings,” *AES preprint 2460 (G-5).*
18. Don and Carolyn Davis, “In-the-Ear Recording and Pinna Acoustic Response Playback,” *AES preprint 2874 (W 3/5-E)*.
19. Joe Ciaudelli, “Large Multi-channel Wireless Microphone Systems,” *Sennheiser Electronic.* (Parts copied with permission.)
20. “RF-Transmission Technique, Wireless Microphones and Reporters Sets,” *Sennheiser Electronic.*
21. G.W. Plice, “Microphone Accessory Shock Mount for Stand or Boom Use,” *J. Audio Eng. Soc.,* Vol. 19, no. 2, February 1971. (Parts copied with permission.)

Bibliography

J. Eargle, *The Microphone Book,* Focal Press, Burlington, MA, 2004.
Microphones, New York: AES, 1979.
Don Davis and Eugene Patronis, Jr., *Sound System Engineering,* Focal Press, Burlington, MA, 2006.
David Miles Huber, *Microphone Manual,* Howard W. Sams & Co., Indianapolis, IN, 1988.

Chapter 17

Loudspeakers

by Jay Mitchell

17.1 Introduction 597
  17.1.1 Uses of Loudspeakers 597
  17.1.2 Loudspeaker Components 597
17.2 Transducer Types 598
  17.2.1 Electrodynamic Transducers 598
  17.2.2 Diaphragm Types 598
  17.2.3 Suspension Methods 599
  17.2.4 Mechanical Construction 600
17.3 Compression Drivers 600
17.4 Electrostatic Transducers 602
17.5 Piezoelectric Loudspeakers 604
17.6 Motor Design Considerations 605
  17.6.1 Output Limitations 606
  17.6.2 Heat Transfer Designs for High-Power Woofers 609
17.7 Radiator Types 609
  17.7.1 Direct Radiators 610
  17.7.2 Cone Radiators 610
  17.7.3 Dome Radiators 610
  17.7.4 Ring Radiators 611
  17.7.5 Panel Radiators 611
  17.7.6 Horns 611
    17.7.6.1 Radial Horns 612
    17.7.6.2 Multicell Horns 612
    17.7.6.3 Controlled Directivity Horns 612
    17.7.6.4 Voice Warning Horns 613
    17.7.6.5 Asymmetric Directivity Horns 613
17.7.6.6 Acoustic Lenses 614
    17.7.6.7 Folded Horns 614
    17.7.6.8 Special Considerations for Low-Frequency Horns 617
17.8 Loudspeaker Systems 617
  17.8.1 Configuration Choices 618
  17.8.2 Types of Loudspeaker Systems 618
  17.8.3 Performance Issues in Multiway Systems 619
  17.8.4 Line Arrays 620
  17.8.5 Crossovers 622
    17.8.5.1 Effect on Maximum Output 623
    17.8.5.2 Effect on Loudspeaker Response 623
    17.8.5.3 Two-Way Crossovers 624
    17.8.5.4 Beyond Two-Way Systems 625

        17.8.5.5 Passive versus Active Crossovers . . . 625
    17.8.6 Acoustic Boundaries . . . 625
    17.8.7 Conclusion . . . 627
17.9 Characterization of Loudspeaker Performance . . . 627
    17.9.1 Motivation . . . 627
    17.9.2 Efficiency and Sensitivity . . . 627
    17.9.3 Network Transfer Function . . . 628
    17.9.4 Loudspeaker Transfer Function . . . 628
    17.9.5 Impedance . . . 629
    17.9.6 Distortion . . . 629
    17.9.7 Characterization for Design Purposes . . . 631
    17.9.8 Characterization for the User . . . 631
17.10 Direct Radiation of Sound . . . 632
    17.10.1 Acoustics of Radiators . . . 633
        17.10.1.1 Piston in an Infinite Baffle . . . 633
        17.10.1.2 Piston Directivity . . . 634
    17.10.2 Direct Radiator Enclosure Design . . . 635
        17.10.2.1 Sealed-Box Systems . . . 635
        17.10.2.2 Vented Boxes . . . 636
        17.10.2.3 Measurement of Thiele-Small Parameters . . . 637
    17.10.3 Horns . . . 638
17.11 Loudspeaker Testing and Measurement . . . 638
    17.11.1 Linear Transfer Function . . . 638
        17.11.1.1 Spectrum Analysis Concepts . . . 638
    17.11.2 Chart Recorders . . . 639
    17.11.3 Real Time Analyzers . . . 639
    17.11.4 Time-Windowed Measurements . . . 639
        17.11.4.1 Measurement of the Impulse Response . . . 640
        17.11.4.2 Maximum Length Sequence Measurements . . . 640
        17.11.4.3 Other FFT-Based Measurements . . . 640
    17.11.5 Swept Sine Measurements . . . 641
Reference . . . 641
Bibliography . . . 641

17.1 Introduction

A loudspeaker is a device that converts electrical energy into acoustic energy (electroacoustic transducer), or more generally, a system consisting of one or more such devices. Loudspeakers are present in our daily lives to such an extent that, in most modern societies, one is in almost constant contact with them. From the time the speaker in our clock radio wakes us in the morning until we turn off the television before we go to bed at night, we encounter loudspeakers almost constantly. Even our computers have loudspeakers.

A general treatment of loudspeakers, including their history and design considerations, in order to fit within a single chapter of a book such as this, is limited to providing an overview of the subject rather than an in-depth treatment of design and theoretical considerations. We will touch on as many of the relevant areas as available space permits, while providing references for the reader who is interested in further study. This chapter may serve as an overview of the subject for end users and audio enthusiasts and as a guide to further study for those interested in performing loudspeaker design work themselves.

17.1.1 Uses of Loudspeakers

Even though there is a very wide range of applications for loudspeakers, they may be thought of as serving some combination of four primary purposes:

1. Communication.
2. Sound reinforcement.
3. Sound production.
4. Sound reproduction.

While there are common requirements for all of these uses, each one also imposes its own demands on loudspeaker attributes. In a given application, it is possible that more than one of these purposes must be served by a single loudspeaker. In such cases, the suitability of the loudspeaker for one or more of its uses may be compromised in order to facilitate others.

Communication. Ranging from intercom systems in offices and schools to radio communications systems for the space shuttle, voice communication systems make our everyday lives safer and more convenient. The first practical loudspeaker was in the earpiece of the original telephone. Since that time, loudspeakers have been an integral part of voice communication systems, from intercom systems to satellite-based telephone and conferencing systems.

Sound Reinforcement. In numerous situations involving public speaking and musical performance before audiences in halls, auditoriums, amphitheaters, and arenas, the sound created by the voices and/or musical instruments is not of sufficient loudness to be heard or understood satisfactorily by everyone present. In such situations, a sound reinforcement system can provide the acoustic gain required to overcome this deficiency.

Sound production. There are a number of subcategories of this type of loudspeaker usage. Perhaps the most readily recognizable is the use of amplification as an integral part of certain musical instruments—e.g., electric guitar, bass, and keyboards. Other examples include emergency warning and sonar systems. Loudspeaker characteristics may be very highly specialized when they are used as part of a sound production system, and loudspeakers optimized for this type of use are often not well suited to other uses.

Sound Reproduction. Playback of recorded music, motion picture soundtracks, and videotape requires a sound reproduction system. Almost every home in the United States has one or more sound reproduction systems. Movie theaters and recording studios also require sound reproduction systems. One of the author's past design projects was a loudspeaker system for use in an international chain of large-screen specialty theaters.

17.1.2 Loudspeaker Components

It is useful to identify the component parts (or subsystems) of a loudspeaker for individual examination and analysis. For purposes of this chapter, the components of a loudspeaker are:

1. Transducer.
2. Radiator.
3. Enclosure.
4. Crossover.

We will examine various forms of each of these components in the sections that follow. Their interactions with each other within a loudspeaker will be discussed. We will also present concepts of loudspeaker performance characterization and an overview of electroacoustic models. The reader is encouraged to pursue the subject matter that is presented here through the references provided in the bibliography. The design and analysis of loudspeakers is a multidisciplinary field, incorporating elements of music, physics, electrical and mechanical engineering, and instrumentation. The individual subject areas are challenging and fascinating in and of themselves, and their convergence in the field of loudspeaker design results in one of the most complex combinations of art and science that has ever existed.

17.2 Transducer Types

There are a number of ways in which electrical energy can be converted into acoustic energy. Of all the possibilities for carrying out this function, a relative few have become dominant in practical loudspeakers: electrodynamic, electrostatic, and piezoelectric. In general, an electroacoustic transducer contains three elements: motor, diaphragm, and suspension. The motor converts electrical energy into mechanical (motional) energy and the diaphragm converts mechanical energy into acoustic energy (vibration of the transmission medium, usually air). A suspension supports the diaphragm, allows it to move in an appropriately constrained fashion, exerts a restoring force proportional to displacement from its equilibrium position, and provides a damping force proportional to the velocity of motion that serves to prevent the diaphragm from oscillating in an undesired manner.

17.2.1 Electrodynamic Transducers

The most common type of transducer used in loudspeakers is the electrodynamic driver. In this type of transducer, a time-varying current passing through a conductive coil suspended in a time-invariant magnetic field creates a force on the coil and the parts to which it is attached. This force causes the parts to vibrate and to radiate sound.

There are a number of viable implementations of electrodynamic transducers. By far the most common is the cone driver. In a cone driver, a cone-shaped diaphragm is suspended at its outer periphery by a structure called a *surround* and (usually) near its center by a *spider*. The *motor* consists of a permanent magnet assembly that concentrates the magnetic field in an annular gap, in which is placed a voice coil that is attached to the center of the cone via a cylindrical coil former. An electrical signal is applied to the voice coil, and the current in the *voice coil* interacts with the magnetic field in the gap to create a time-varying force that vibrates the diaphragm. Fig. 17-1 shows a typical cone driver. The most commonly used magnetic material is ferrite, or ceramic. Other magnetic materials used in loudspeakers include aluminum/nickel/cobalt (*alnico*) and neodymium/iron/boron, (*neodymium* or *neo*). The magnet structure is typically held together with an anaerobic thermoset adhesive. Some loudspeakers are assembled with bolts through the magnet. In this case, stainless steel or brass screws must be used, so as not to magnetically short the top plate to the back plate. A rear cover may or may not be used. A vent through the pole piece may be provided. It serves to prevent the addition of a spring constant due to the small air cavity under the center cap (dust cover) and to reduce turbulence-induced noise due to pumping effects in the magnet gap.

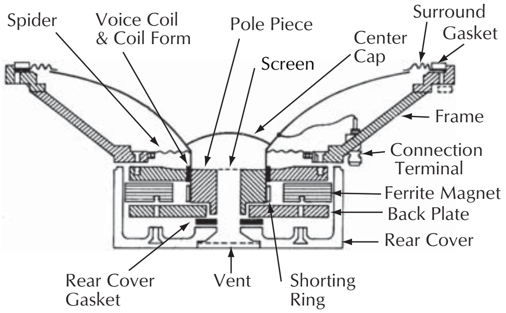


Figure 17-1. Typical woofer parts identification. Courtesy Yamaha International Corp.

17.2.2 Diaphragm Types

The most common direct-radiation device is the cylindrical voice coil–driven paper cone. The cheapest cone to make is the folded cone, which is cut from a sheet of paper, rolled, and bonded at the seam. A more expensive and difficult to make cone is the molded-paper cone. These are one piece, molded by straining a slurry of water and paper pulp through a strainer mold in the shape of the desired end product. The formed wet mat of pulp is then pressed and baked to remove residual moisture, bearing a dry, strong one-piece cone, free of joints. Ribs and concentric rings are sometimes molded into the cone, and the cones can be formed with straight or curved sides of varying depth. These are all available from suppliers of cones.

While most mathematical models of a direct radiator assume a rigid piston, in practice this is impossible to achieve. In some cases, diaphragm rigidity is intentionally reduced in order to produce specific desired behavior. Two examples involving a controlled breakup are shown in Figs. 17-2 and 17-3. The whizzer cone in Fig. 17-2 is intended to radiate high frequencies as the larger cone decouples from the motor. The Biflex principle, as popularized by Altec Lansing in the 1950s, is shown in Fig. 17-3. The inner cone is attached via a compliant element at A to the large outer cone in hopes of decoupling the outer cone at high frequencies.

Damping dope is applied to the coupling connection in an attempt to smooth the decoupling transition frequency response. While whizzer cones are still in use in some inexpensive ceiling speakers, there are at this time no devices similar to the Biflex on the market.

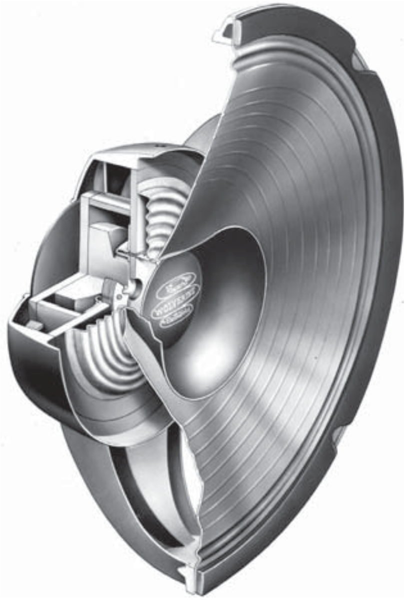


Figure 17-2. Loudspeaker incorporating a whizzer cone.

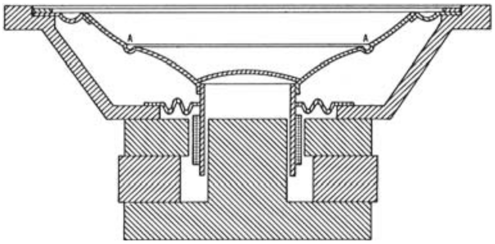


Figure 17-3. Loudspeaker illustrating decoupling center cone. (From U.S. Patent 4,146,756.)

In addition to felted paper, a number of newer materials have found use in cone-type low- and medium-frequency loudspeakers. A variety of plastics have been used, the most popular being polypropylene and bextrene. The KEF Company introduced a composite aluminum-skinned foam-core sandwich cone. Community Professional Loudspeaker's M4 compression midrange similarly uses a carbon fiber/epoxy composite diaphragm. Adamson Acoustics in Canada uses a Kevlar fabric resin-bonded diaphragm for the midrange driver. Mitsubishi Electric (Japan) introduced a studio monitor, which used cone woofers fabricated from a honeycomb core/carbon fiber skin composite.

In a loudspeaker with an alnico magnet, the magnet is directly under the pole piece (as opposed to being between the top and backplates), and the outside of the magnet structure is a cast iron return from the bottom of the magnet to the top plate. Venting may be accomplished via a hole covered with open wire mesh in the center dome. Other methods include a uniformly porous dome with no magnet vent, Fig. 17-4.

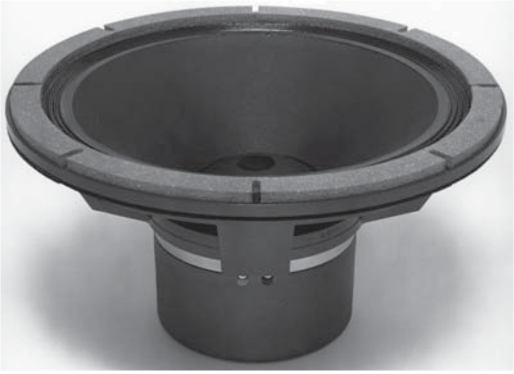


Figure 17-4. Alnico magnet woofer—Altec 515-8LF. Courtesy of Altec Lansing Corp.

17.2.3 Suspension Methods

The suspension of a cone driver comprises two distinct components: the surround and the spider. The surround is attached to the periphery of the diaphragm or cone and is itself attached to the support structure (the basket in the case of a cone driver). The spider is attached to the voice coil former (or to the cone in the vicinity of the former) and is also attached to the basket on its periphery. Because they affect cabinet sealing, surrounds are designed to be nonporous. Surrounds and spiders both contribute to the damping of the motion of the diaphragm. The most popular surround construction is heat-formed, open-weave, resin-impregnated linen with formed-in convolutions and sealed with damping dope. Other surrounds are made of foam or butyl rubber formed in a half-roll. On some loudspeakers, a viscoelastic (never-drying) dope is applied to the surround.

Spiders are usually made of a heat-formed, open-weave, resin-impregnated cloth that is formed into convolutions. They are usually not treated with a sealing material (dope). The unsealed fabric is needed for venting, since the air beneath the spiders can otherwise be trapped. This also tends to damp the spider. An early method of making porous spiders was to die-cut them from solid phenolic-impregnated linen sheet stock. The spider is not required to seal the edge of the cone to its enclosure as is the surround. In a typical cone driver, the spider contributes the majority of the stiffness in the suspension.

17.2.4 Mechanical Construction

The Peavey Black Widow bass drivers are unusual in that they have a streamlined magnet structure, called focused-field geometry, Fig. 17-5. It employs a magnetic circuit that has smoothly flowing flux lines, as might be intuitively preferred for a fluid flow channel. Other manufacturers have adopted similar approaches to magnet design. An added benefit of this approach is minimization of weight.

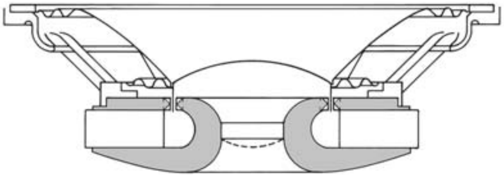
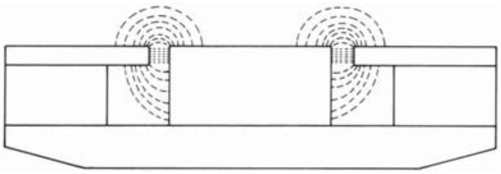
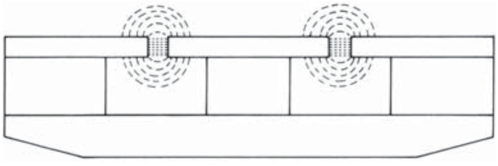


Figure 17-5. Peavey focused-field geometry magnet structure with one-piece backplate/pole piece forging. Courtesy Peavey Electronic Corp.

JBL ferrite-magnet drivers have symmetric field geometry. Fig. 17-6 shows the top plate configuration, which makes the magnetic leakage flux at the top and bottom of the gap symmetric, thereby, according to the manufacturer, reducing magnetic drive asymmetry and the resulting low-frequency distortion.



A. Flux distribution in nonsymmetrical gap showing an uneven fringe field.



B. Flux distribution with symmetrical field geometry showing equal fringe field on both sides of the gap.

Figure 17-6. JBL symmetric field geometry versus asymmetric design. Courtesy JBL.

Another form of electrodynamic transducer is the dome radiator. Most commonly used for high frequencies, dome radiators have the advantages of compactness and predictability of acoustic behavior. Domes can be made from linen, impregnated phenolic fabric, Mylar™, paper, aluminum, titanium, beryllium, and composites such as carbon fiber/epoxy. Soft dome tweeters have been in widespread use for a number of years. Some of this popularity may be due to the fact that there is no abrupt transition from piston radiation to breakup. Instead, most of the radiation from a soft dome comes from the region immediately adjacent to the voice coil, making it function as a ring radiator.

Several flexible diaphragms have been used on magnetic drivers, all sharing the same basic construction: etched aluminum conductors on Mylar™ film. These are operated in various magnetic field configurations to produce sound. One of the earliest of these is the Magneplanar® loudspeaker, which consisted of an entire field of magnets over which the diaphragm conductor was mounted. Magneplanars are in the shape of large panels. The Heil high-frequency driver, used in systems manufactured by ESS, used direct radiators similar to Magneplanar® in that the voice coil was printed on Mylar™. The ESS-Heil unit, however, was corrugated, and the sound was produced by these vertical pleats moving open and closed, thereby squeezing air into radiated sound. An extension of this was used also by ESS in the Transar system, which used hollow spheres modulated by electromagnetically driven rods. Mitsubishi Electric (Japan) developed a printed conductor high-frequency device called the *leaf tweeter*, as shown in Fig. 17-7. The ribbon loudspeaker is the simplest and has excellent potential for good high-frequency response due to the fact that the diaphragm is the conductor. No extra diaphragm structure is used on the ribbon.

17.3 Compression Drivers

One means of improving the performance of an electrodynamic transducer that will be used to drive a horn is to create a compression driver. In a compression driver, the diaphragm radiates into a compression chamber and its output is typically directed through a phasing plug to the driver's exit, which is attached to the throat of the horn.

The advantage of a compression driver is that relatively small diaphragm velocities are converted to larger particle velocities at the exit of the driver. The effect of this transformation is that less diaphragm excursion is required for a given acoustic power output. The tradeoffs

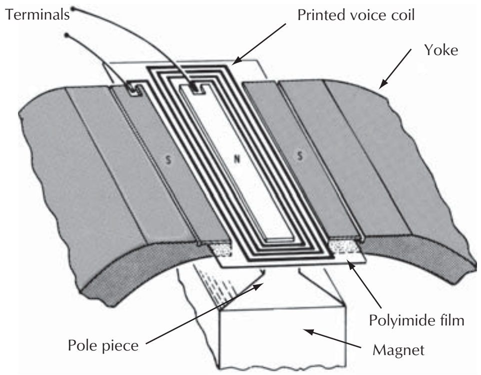

Terminals
Printed voice coil
Yoke
S
N
S
Polyimide film
Pole piece
Magnet

Figure 17-7. Technics Leaf Tweeter diaphragm detail. Courtesy Panasonic Industrial Corp.

for this coupling include possible increases in certain distortion components and the requirement for a horn. Compression drivers are not used as direct radiators.

The purpose of the phasing plug is to equalize path lengths from the diaphragm surface to the exit. To the extent that this is accomplished, the useful bandwidth of the driver will be extended upward in frequency.

Fig. 17-8 is a cross-sectional view of a typical ceramic magnet wide-range compression driver using a dome diaphragm. The case construction is unusual and peculiar to this design by Yamaha. The phase plug is also a bit unusual; however, it is still of the circumferential-slit variety on the phase plug (dome) surface. The diaphragm is aluminum and is supported by a typical Bakelite™ or plastic support frame. The back cap has sound-absorbing material inside to discourage interfering air resonances in the cap.

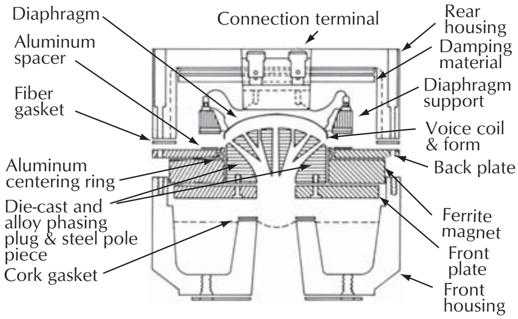

Diaphragm
Aluminum spacer
Fiber gasket
Aluminum centering ring
Die-cast and alloy phasing plug & steel pole piece
Cork gasket
Connection terminal
Rear housing
Damping material
Diaphragm support
Voice coil & form
Back plate
Ferrite magnet
Front plate
Front housing

Figure 17-8. Typical ceramic-magnet compression driver. Courtesy Yamaha International Corp.

Fig. 17-9 shows a 2 inch throat JBL driver, model 2440, using an alnico magnet. The phase plug is more typical than that in Fig. 17-8, using more straight through circumferential slits. The JBL plug is made of cast Bakelite.

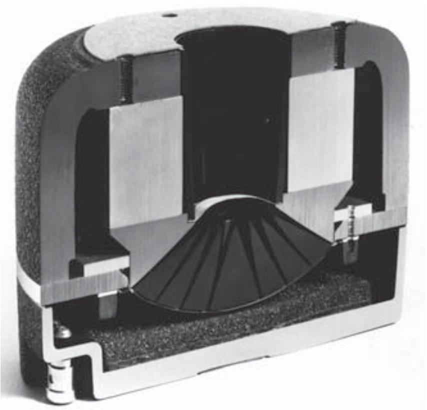


Figure 17-9. JBL 2440 2 inch throat alnico compression driver. Courtesy JBL/UREI.

Fig. 17-10 shows another alnico driver, the 1 inch Altec 802/808. The 802 uses an all-aluminum diaphragm with tangential surround coupled via the phase plug and expanding throat section to a 1 inch diameter exit. The 808 is identical to the 802 in all respects except for the diaphragm. From left to right in the exploded view are the pot, the alnico magnet slug that fits under the pole piece, in which is mounted on a radial-slit tangerine phase plug. This unusual design is made from glass fiber–filled plastic and is bonded to the pole piece. Above this (to the right) is the ring that centers the pole piece in the air gap via the top plate. It is nonmagnetic (brass), and the holes provide a mechanical load on the voice coil, which affects response and distortion. Next are the diaphragm assembly and rear cap.

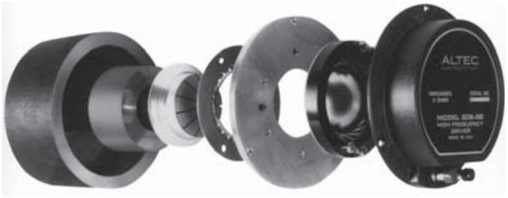


Figure 17-10. Altec 802/808 alnico 1 inch throat compression driver with Tangerine radial phase plug. Courtesy Altec Lansing Corp.

The preceding three examples are representative of dome diaphragm compression driver design practices, both in concept and in practical implementation.

A large number of variations exist in the art including a wide variety of suspension shapes and materials. Domes are made from aluminum, titanium, and beryllium. Yamaha International Corp. makes its suspension out of stamped beryllium-copper cantilever fingers, and JBL and Ramsa stiffen their suspensions with rhombic and diamond patterns, respectively; this is actually a redistribution of diaphragm breakup resonances.

Midrange compression drivers are useful where there is a requirement to supply very high levels of acoustic power with low distortion. Community Professional Loudspeakers, M4 midrange driver is shown in Figs. 17-11 and 17-12. It is intended for use from 200 Hz to 2000 Hz. The diaphragm is approximately 7 inches in diameter. Originally it was fabricated from specially formed aluminum skins and a light, stiff foam core, about 0.090 inch thick. More recent versions had diaphragms made of a carbon fiber composite.

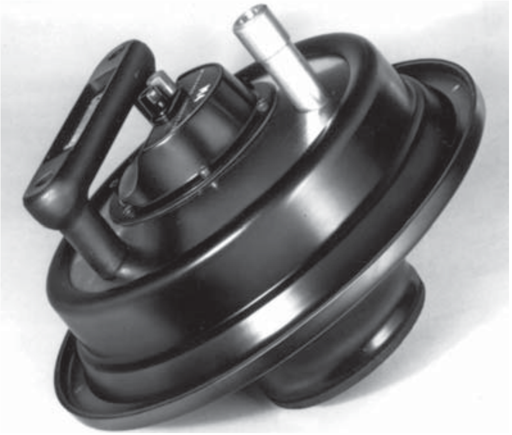


Figure 17-11. Community M4 4 inch throat midrange driver. Courtesy Community Professional Loudspeakers.

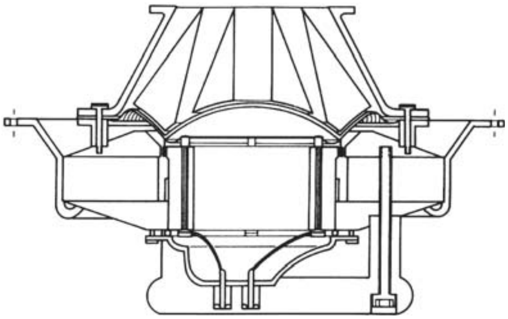


Figure 17-12. Community M4 cross-section view. Courtesy Community Professional Loudspeakers.

Another compression driver configuration is the screw-on driver. The University 7110XC (explosion proof) is shown in Fig. 17-13. This type of unit is often used on a reentrant horn in public address systems. Throat diameter is usually ¾ inch. Diaphragms are most often made of phenolic resin-impregnated domes with integral convoluted suspensions. Voice coils are usually round copper wire.

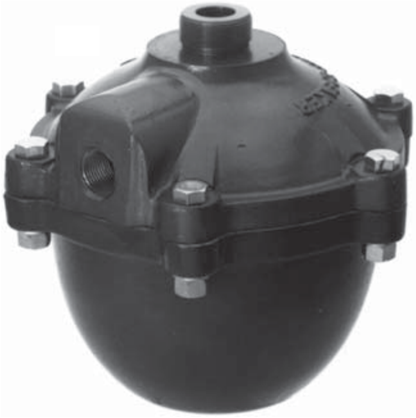


Figure 17-13. University 7110XC ¾ inch throat public address driver.

17.4 Electrostatic Transducers

Electrostatic transducers make use of the fact that two static electrical charges placed at a distance from each other will experience a force directed along a line between them. The force is attractive if the charges have the opposite sign (positive and negative) and repulsive if the charges have the same sign (positive and positive or negative and negative). In practical loudspeaker designs, the forces are attractive, due to the complementary nature of the charge transfer from the amplifier output to the speaker plates. The magnitude of the force is inversely proportional to the distance between the charges and directly proportional to the magnitude of the charges.

A typical electrostatic loudspeaker consists of a diaphragm made of two pieces of metallic foil separated by a sheet of dielectric, or nonconductive, material. By itself, the application of a pure ac signal (i.e., one with no dc component) to an electrostatic loudspeaker would cause attractive forces for both positive- and negative-going signal excursions, since the induced charges are opposite in both cases. This would create a frequency-doubled signal containing extremely high levels of harmonic distortion.

For this reason, a dc polarizing voltage is applied to the foil diaphragms, maintaining a steady attraction between them. The audio (ac) signal is superimposed on this dc offset, modulating the attractive force. In response to this modulated force, the diaphragms to move opposite (toward or away from) each other. The upper limit on the amplitude of the allowed signal voltage is then equal to half the polarization voltage. This arrangement is the basis of all modern electrostatic loudspeakers. The result is an acceptably low level of harmonic distortion, as long as variations in the distance between plates or the diaphragms are minimized. The movement of the foil diaphragms generates sound waves. The diaphragms produce equal acoustical power radiated in opposite directions. This set of characteristics defines a dipole radiator.

It is asserted by the designers of electrostatic loudspeakers that they overcome certain basic disadvantages of cone-type loudspeakers, particularly with respect to the propagation of acoustic energy at the high frequencies. Cone-type loudspeakers are driven by a voice coil that is attached to a relatively small portion of the total diaphragm area, and they do not behave as pistons at higher frequencies. Because the electrostatic loudspeaker has a diaphragm that is driven uniformly across its surface, breakup is said to be eliminated. Additionally, the diaphragm can have low mass compared to the air load on the diaphragm. This enhances high-frequency and transient response.

Electrostatic loudspeakers may be constructed in several different ways. Two of the most common construction types are:

1. Stretching the diaphragm between supports around its periphery and leaving an air gap between the diaphragm and two stationary electrodes, Fig. 17-14.
2. Using an inert diaphragm that is supported by a large number of tiny elements disposed across the entire surfaces of the two electrodes. These elements act as spacers to hold the diaphragm in the center between the electrodes, Fig. 17-15.

In the latter type of loudspeaker, the diaphragm is a thin sheet of plastic on which has been deposited a very thin layer of conductive material. It is supported by multiple small elastic elements that hold the diaphragm in place but permit it to follow audio-signal waveforms. The electrodes on each side of the diaphragm are acoustically transparent to avoid pressure effects from trapped air as well as to permit acoustic energy to propagate away from the diaphragm. This type of construction permits the diaphragm to be of arbitrary size. The performance per unit area is the same for any area of the diaphragm. The actual loudspeaker is a thin surface curved in the horizontal, forming a section of a cylinder.

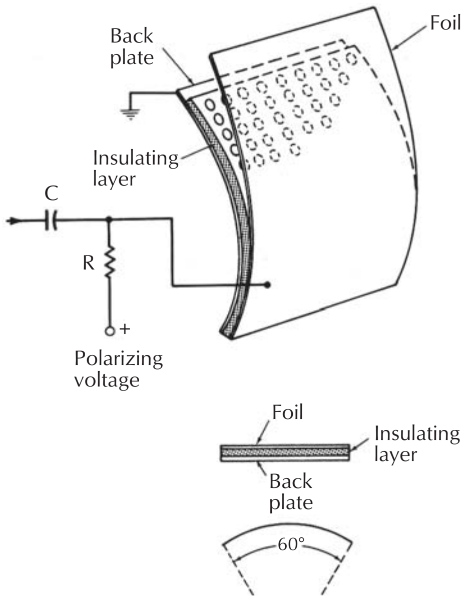


Figure 17-14. Electrostatic- or capacitor-type loudspeaker.

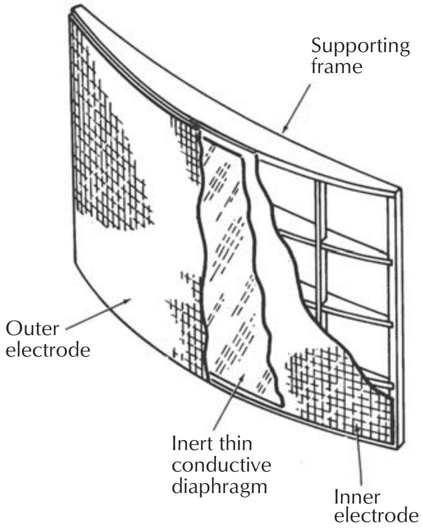


Figure 17-15. Cutaway view showing the internal construction of an electrostatic loudspeaker.

A surface that is large with respect to wavelength becomes increasingly directional at high frequencies.

Since an electrostatic loudspeaker is designed to couple directly with the acoustic resistance of air, the mass of the diaphragm is quite small and can be neglected with little effect on the accuracy of predictive models. The velocity of the diaphragm is directly proportional to the electrostatic force applied, except as altered by the stiffness of the diaphragm suspension. Measurements indicate that for a constant voltage applied to the electrodes, the acoustic response is uniform (flat) to well beyond the range of human hearing.

Output at low frequencies is limited by the maximum linear amplitude of the diaphragm motion, which is determined by spacing between the diaphragms and damping in the suspension. The maximum power output from an electrostatic loudspeaker of a given diaphragm area is determined by the strength of the electrostatic field that can be produced between the diaphragm and the electrodes.

An electrostatic loudspeaker is seen by an amplifier as a capacitor with a value on the order of 0.0025 μF from electrode to electrode. Thus, the magnitude of the impedance presented by the loudspeaker to the output of the amplifier falls off at 6 dB per octave as the frequency is increased. This presents some problems for driving electrostatics, as many amplifiers are not designed to drive purely capacitive loads.

Because electrostatic loudspeakers are relatively large in area compared to cones, their directivity is high in comparison to cone systems. Various schemes have been used by designers of electrostatics to address this issue. The Quad ESL63 is one example. Here, the diaphragm is broken into different regions for different frequency ranges, the smaller ones being used for higher frequencies, thereby making them wider in dispersion than a large single panel.

17.5 Piezoelectric Loudspeakers

Piezoelectricity, or pressure electricity, was discovered in the 1880s by the Curies. It is today a feasible motor drive mechanism for loudspeakers. In a piezoelectric material, a voltage applied to the material will result in a mechanical strain or deflection. The reverse is also true, and piezoelectric elements can be used in microphones. This characteristic is attractive for direct-drive units such as ultrasonic devices. For loudspeakers, however, some means must be applied to mechanically amplify the inherently low excursion so that a loudspeaker diaphragm may be driven properly.

One of the earliest discovered piezoelectric substances is Rochelle salt. Although Rochelle salt is still widely used, it suffers from poor mechanical strength, low temperature breakdown (55°C), and extreme sensitivity to humidity. Barium titanate is the first piezoceramic to be developed. Although it is not as electrically sensitive, it is still widely used, exhibiting many superior characteristics over Rochelle salt. The most widely used piezo material today is lead zirconate titanate, developed first in Japan in the 1950s. This material (PZT) is now highly refined and exhibits the best properties of any piezo material for loudspeaker use.

PZT material is formed by baking a ceramic slurry or clay into bars about 1 inch in diameter and then slicing the bars into thin wafers. Two wafers are bonded together in opposing polarity, with electrodes on their flat surfaces, forming a bimorph bender. As voltage is applied to the bender, deformation of the disc results in greater displacement at its center.

Early commercial attempts at the application of bimorph benders to loudspeaker cones involved a rectangular drive element anchored at three corners, allowing the fourth corner to drive the loudspeaker cone fore and aft. Other attempts used a cantilever structure anchored at one end with the loudspeaker cone mounted at the other. In 1965 when Motorola, Inc. first manufactured a piezoelectric loudspeaker, they used a length expander tube driving a horn-loaded cone directly. This device, like most piezoelectric loudspeakers made until that time, still lacked sufficient voltage sensitivity to be coupled directly to conventional systems without using an auxiliary step-up transformer.

The development of the circular bimorph using a corrugated center vane represented the next step forward in piezo loudspeaker technology. The action of the two disks working against each other, one expanding while the other contracts, functions as a mechanical transformer, giving impedance reductions of about 20:1. The basic operation is as follows: The driver dishes in and out; it pumps the cone fore and aft or, in the case of the horn, into a compression chamber that is then coupled to the throat of a horn via a slot and flared rib construction. The driver is allowed to hang free in space, working against its own inertia to pump the cone.

Further advancements in the state of the piezoelectric art came from Tamura and coworkers in their work on piezoelectric high-polymer films. This concept of a diaphragm possessing piezoelectric properties and thus coupling directly to the air without the use of any separate motor structure represents a substantial advancement toward the ideal acoustic transducer.

Another problem area in the development of the PZT loudspeaker was in the power-handling capability of the driver. The theoretical failure mode of a piezoelectric tweeter is the depolarizing of the driver through excessive drive level and/or high temperature. The Curie point (depolarizing temperature) of the PZT used here is above 150°C, and the depolarizing voltage is 10 V/25 μm thickness, or about 35 Vrms for the basic driver. These numbers describe a fairly impressive power-handling capability, but unfortunately one that is

if

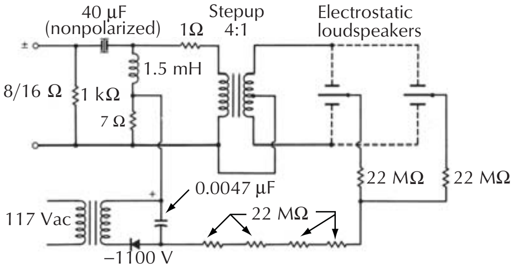


Figure 17-16. Typical coupling circuit and high-voltage power supply for electrostatic loudspeakers.

reached only in theory. In reality, under continuous high drive levels, mechanical stress on the surface of the ceramic wafers generates cracks in the microstructure that eventually penetrate the entire wafer. This is especially severe around the area where solder connections are made to the wafers, since the soldering operation tends to prestress the material at this point. The net result is that the 35 V maximum drive level is an intermittent specification, with the continuous drive level recommended at a 15 V maximum up to 20 kHz. For use above that frequency, it has been recommended that the level be reduced further through the addition of a series attenuation resistor so as to safeguard the ceramic element from absorbing excessive high-frequency power. Here again, when using larger, thinner ceramic wafers, these problems are further aggravated. Using this ceramic is an area for future development.

Fig. 17-17 shows the Motorola KSN 1001A. Although Motorola manufactures a wide variety of other piezo-driven loudspeakers, the one illustrated here is the most widely used.

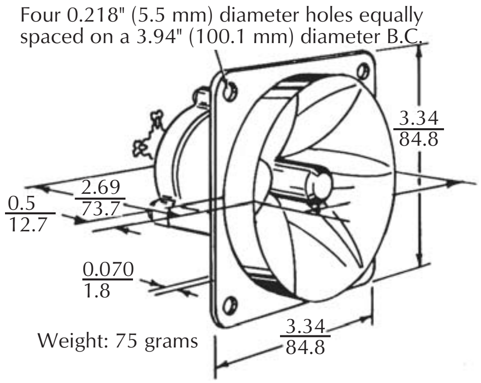

Four 0.218" (5.5 mm) diameter holes equally spaced on a 3.94" (100.1 mm) diameter B.C.
3.34
84.8
0.5
12.7
2.69
73.7
0.070
1.8
Weight: 75 grams
3.34
84.8

Figure 17-17. Motorola KSN 1001A piezoelectric ultra-high-frequency driver/horn. Courtesy Motorola, Inc.

One near-optimal application of piezoelectric drive is underwater use. This is due to the excellent impedance match of the piezoelectric material to water via a waterproof barrier. Lubell Labs manufactures the underwater loudspeaker shown in Fig. 17-18. Although swimming pool loudspeakers using standard electromagnetic drivers are also available, the piezoelectric configuration is more efficient due to its mechanical impedance match to water. The loudspeaker is fixed to the side of the pool and driven like a conventional loudspeaker. Lubell Labs makes high-power arrays of these devices and a portable swim coach system with a noise-canceling microphone for underwater communications in various pool athletic events.

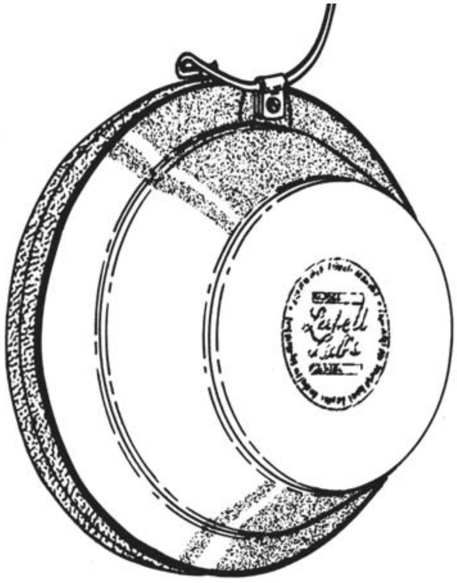


Figure 17-18. Lubell Labs underwater piezoelectric loudspeaker. Courtesy Lubell Laboratories, Inc.

17.6 Motor Design Considerations

The most common means of coupling amplifier output to the diaphragm in an electrodynamic transducer is via a cylindrical voice coil. This configuration is used on all magnetic cone loudspeakers and compression drivers. This is commonly known as a linear motor. The coil, made of round or rectangular wire (edgewound), is wound around a hollow cylinder called a *former*. Formers may be made of paper, plastic (e.g., Kapton polymer, Mylar™), or aluminum. The voice coil assembly is bonded to the diaphragm. Fig. 17-19 shows the construction of a typical cone loudspeaker.

One novel approach to motor design involves printing or etching a conductor onto a thin sheet of Mylar™ (0.0005 inch) then folding it to produce and

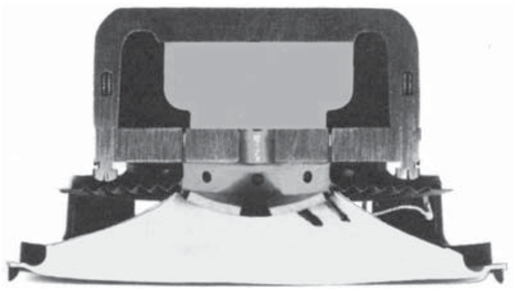


Figure 17-19. Cross section of a typical cone loudspeaker showing construction (alnico magnet at center under pole piece). Courtesy JBL.

pleated diaphragm that is forced in the magnetic field. In another implementation, continuous lengths of wire are bonded to a large panel of Mylar™, which is operated over a field of bar magnets. The leaf tweeter is similar, etching a conductor field on Mylar™. They are identical in principle to Fig. 17-19 and are discussed more thoroughly elsewhere in this text. The ribbon loudspeaker, Fig. 17-20, is a special case in which the voice coil serves as both conductor and diaphragm.

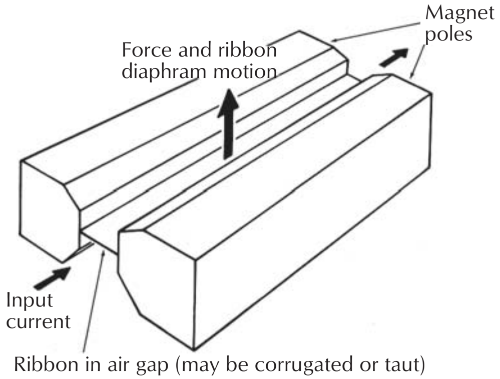

Magnet poles
Force and ribbon diaphram motion
Input current
Ribbon in air gap (may be corrugated or taut)

Figure 17-20. Ribbon loudspeaker.

One notable departure from conventional linear motor design is the Servo-Drive loudspeaker. This patented drive system uses a rotary servomotor that drives a woofer cone and suspension assemblies via a pulley-belt mechanism, alternately pushing and pulling the diaphragms in response to the input signal. Two opposing diaphragms are driven in a push-pull arrangement so as to yield a balanced axial force on the drive mechanism. The motor is configured so that it presents a typical impedance load to an amplifier. SDL (for servo-drive loudspeakers) speakers come in a variety of sizes and power capacities, but typically they are in the form of low-frequency horns. Fig. 17-21 shows the mechanism employed to translate rotational motion of the servomotor to the linear motion needed to drive the opposing diaphragm assemblies. The opposing reinforced blastomeric belt mechanisms are used in the rotation-to-linear conversion, and the result is noiseless and free of slip. The opposing diaphragms drive the throats of conventional wood-fabricated folded bass horns. The positions of the diaphragms are shown in Fig. 17-22. Fig. 17-23 shows the positioning of the servo-driven diaphragms in a typical folded bass horn.

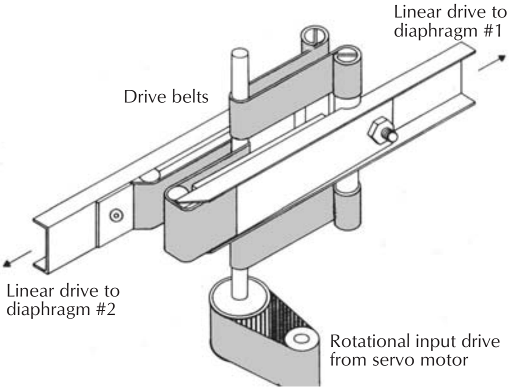

Linear drive to diaphragm #1
Drive belts
Linear drive to diaphragm #2
Rotational input drive from servo motor

Figure 17-21. Belt drive system of the SDL loudspeaker.

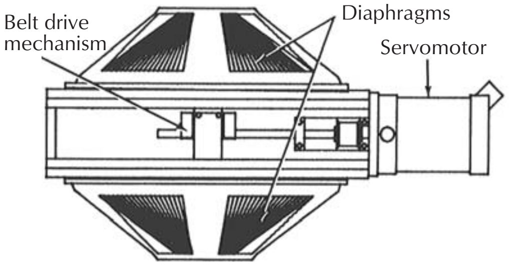


Figure 17-22. Position of SDL belt drive and opposing diaphragms.

17.6.1 Output Limitations

The maximum usable output of an electromagnetic loudspeaker is a function of a number of parameters, including diaphragm displacement, heat transfer, sound quality (maximum acceptable nonlinearity), and/or wear life due to fatigue of moving parts.

There are two fundamental limitations on a magnetic driver, a displacement limit and a thermal limit. Displacement limits may be caused by either mechanical or electrical factors. Mechanical displacement limiting occurs when a moving part contacts a

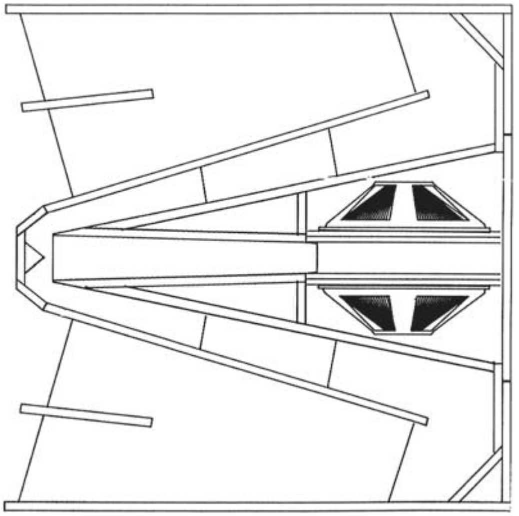
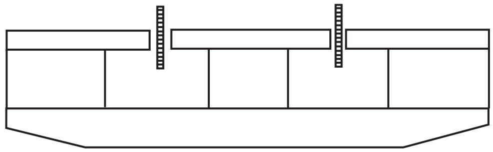
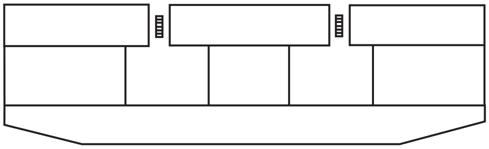


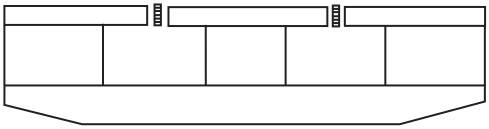
Figure 17-23. Position of the SDL diaphragms on a folded horn.



A. Overhung coil.



B. Underhung coil.



C. Equal length coil and gap.

Figure 17-24. Three basic voice coil/magnetic gap configurations.

stationary one or when a suspension element is made unacceptably nonlinear (either temporarily or permanently) by deformation beyond its design range. Electrical displacement limiting occurs when the motor is operated outside its range of linear travel. This is a function of the length of the windings on the voice coil and the thickness of the plates that form the magnet gap. Fig. 17-24 shows three typical voice coil configurations: equal length, overhung, and underhung coils. When any of these coils reaches a displacement that causes a reduction in the current sensitivity of the motor, higher distortion will result.

It has been empirically determined that, due to a magnetic fringe or leakage field at the pole tips, an excursion of 15% farther than the gap length results in a reasonable distortion level (approximately 3% harmonic distortion at low frequencies). The equal length voice coil, Fig. 17-24C, has the greatest potential for motor-generated distortion. However, it also yields the highest motor strength (the greatest total conductive mass in the highest density magnetic field). The equal length voice coil is a common configuration for compression drivers, where maximum excursion is intrinsically low. The underhung coil, Fig.17-24B, allows greater excursion but requires a larger magnet due to the longer gap. For moderate flux density levels (10,000 to 15,000 G), this design, as compared to the equal length design, requires approximately twice the magnet weight (twice the area and the same length) for a doubling of the gap length. This approximately doubles the excursion capacity, giving four times the acoustic power output capability (6 dB) for a doubling of magnetic weight (3 dB). The overhung coil, Fig. 17-24A, is capable of the greatest motor linearity, all else being equal. It is commonly seen on woofers used as direct radiators, where higher excursion is required. The major disadvantage here is that the coil that is not in the gap does not participate in transduction. The extra coil length does add both mass and dc resistance, however, reducing motor efficiency. In spite of this, there are numerous examples of successful commercial woofers using overhung coils. The transducer designer must take into account the often conflicting demands of high-efficiency, high-output, and low-frequency extension to arrive at an optimum design for a given range of applications.

The thermal limit of a magnetic loudspeaker motor is a function of the temperature limits of the materials used and heat transfer from the coil assembly to the outside world. Most adhesives used in the loudspeaker industry have an upper limit between 120°C and 177°C (250°F and 350°F). Some epoxy adhesives will tolerate higher temperatures, but they can require special curing processes and are therefore potentially more difficult to use. Wiring insulation may tolerate temperatures as high as 218°C (425°F). Anodized aluminum wire has the melting point of aluminum as a limit. Voice coils operated at high temperatures have higher resistance. A 1°C rise produces approximately a 0.4% rise in dc resistance in both copper and aluminum. Therefore, operating a voice coil 100°C above ambient (127°C or 261°F) will cause the voice coil resistance to increase to 40% above its ambient value. The following equation give voice coil resistance at any temperature in degrees Celsius

$$R_T = R_o + 0.004(T - T_o) \tag{17-1}$$

where,

R_T is the resistance at temperature T in ohms,

R_o is the resistance at ambient temperature T_o in ohms,

T and T_o have units of °C.

The operating temperature of the voice coil, T_{VC}, is determined by ambient temperature, the amount of power being dissipated in the coil, and a parameter called *thermal resistance*, expressed in degrees Celsius per watt, °C/W. The thermal resistance is a measure of the ability of an object to transfer heat away from itself. The lower the value of the thermal resistance, the more effective the object is at this transfer. As power is doubled, final temperature rise above ambient is doubled. Heat transfer in a loudspeaker is a function of the air gap design, voice coil design, and the ability of the loudspeaker frame and magnet to dissipate heat to the surrounding or ambient air. Referring to Fig. 17-25, the thermal rise T_{VC} of a stationary voice coil in an air gap is

$$\begin{aligned}\Delta T_{VC} &= T_{VC} - T_s \\ &= \frac{QL}{A_T K}\end{aligned} \tag{17-2}$$

where,

T_{VC} is the temperature of the voice coil in °C,

T_s is the temperature of structure (magnet) in °C,

Q is the electrical heating power (IR) in watts,

L is the effective air gap length in inches,

A_T is the total gap area in square inches exposed to the voice coil,

K is the conductivity of air or 7 × 10–4 W/°C.

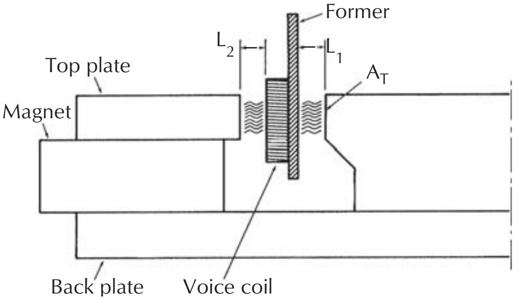


Figure 17-25. Heat conduction in magnetic loudspeakers.

As the air gap length is decreased and the area increased, heat transfer increases (or, equivalently, thermal resistance decreases). Making the voice coil former of aluminum will increase effective heat transfer area; the thicker the aluminum, the greater the effect. Voice coils wound on aluminum formers with large diameters in magnets with large gap areas and very tight coil to gap tolerances are capable of handling high electrical power due to good heat transfer in the air gap. In short, large, accurately constructed loudspeakers can usually handle more power. As the loudspeaker moves, it may be able to pump the air in the gap to improve heat. The loudspeaker designer may be able to exploit this behavior. Given voice coils of the same length, the underhung and equal-length configurations will have greater heat transfer capacity. The overhung coil would only conduct heat well in the gap region, while the coil ends remaining out of the gap would be more likely to suffer damage at high power level because of relatively poor heat transfer. Typical thermal behavior for most coils is on the order of 0.5°C/W to 3°C/W input.

A heat-conducting magnetic liquid may be used to improve heat transfer. Known as *ferrofluids*, these fluids will be retained in a magnetic air gap due to magnetic attraction. Their thermal conductivity is seven to ten times higher than that of air. Since ferrofluid alters the mechanical damping of the moving assembly, its use has implications for the design of the motor assembly. There are also issues related to compatibility of ferrofluid with adhesives and materials used in the construction of a transducer. For these reasons, ferrofluids should generally be designed into a loudspeaker, rather than added on.

Temperature rise in voice coils is not instantaneous. It is directly related to mass. As one might suspect, light voice coils have short thermal rise times, and vice versa. The thermal time constant of a loudspeaker coil (the time required for the coil to reach 63% of its final value) is given by:

$$\begin{aligned}t &= MC\frac{\Delta T}{Q} \\ &= MC\frac{L}{A_T K}\end{aligned} \tag{17-3}$$

where,

t is the time constant in seconds,

M is the mass of the coil,

C is the specific heat of voice coil material in joules per gram in degrees Celsius,

$\Delta T/Q$ is the thermal resistance in degrees Celsius per watt.

For example, a typical copper woofer voice coil has a mass of 24 g and a gap heat transfer coefficient (thermal resistance) of 1°C/W. Copper has a specific heat of 0.092 cal/g°C or 0.0220 J/g°C. Therefore, using Eq. 17-3, $t = 0.528$ s. This is a typical voice coil response time. An aluminum coil will typically have a shorter thermal time constant.

The time constant of the magnetic structure and frame can be on the order of hours. For this reason, long duration power tests are required to evaluate the maximum power tolerance of transducers. Initially, the voice coil might be at 280°F (137°C), but over the course of 2 hours, the mechanical structure (typical 1 to 3°C/W) could rise another 200°F to 300°F (100°C to 150°C), bringing the voice coil well over the thermal limit of its materials and adhesives. Heat transfer from the frame and magnet to the air is another important consideration. Although the rise time is large, the final temperature may vary greatly due to the enclosure. A vented enclosure with vents at the top and bottom with no fiberglass insulation might provide adequate ventilation for a hot loudspeaker. The same loudspeaker in a closed box stuffed with fiberglass might be subject to a dangerously high temperature rise. Attention to this final thermal path is warranted in applications that will demand maximum output from enclosed loudspeakers.

The efficiency of a loudspeaker has a direct bearing on the thermal load it must withstand for a given acoustic output level. The more efficient the loudspeaker, the lower the self-heating for a given output level; all else being equal, a loudspeaker with 3 dB higher overall sensitivity for a given impedance will experience one-half the thermal load for a desired output level.

In concert touring use, loudspeakers are routinely operated at and even beyond their design limits. Given that a loudspeaker that operates at twice its voice coil resistance due to heating will be 6 dB less sensitive, sound quality can vary greatly over the course of a performance. In failure situations, the nature of the input signal will usually determine the type of failure mode. Thermal failure can be precipitated by compressed high-frequency content material (low dynamic range). Mechanical failure is often due to dynamic, percussive material, such as might occur in a recording studio with drum channels set to solo, as well as other signals that do not limit dynamic range. Another cause of mechanical failure, most often in high-frequency transducers, is the application of a highly clipped signal that has been passed through a high-pass filter. Such a signal will contain a peak-to-peak voltage that is twice that of the input signal. This phenomenon is illustrated in the section on crossovers.

17.6.2 Heat Transfer Designs for High-Power Woofers

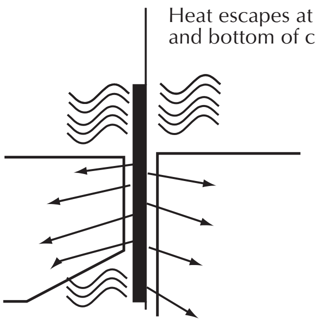
Of all the components in a sound reinforcement system, more heat is generated in low-frequency devices than in any other. While high-frequency horn driver combinations deliver 110–117 dB/1 W/1 m and midrange devices deliver 100–110 dB, woofers rarely exceed 100 dB. A typical woofer in a vented enclosure is in the 94–97 dB/1 W/1 m range. These devices are typically 2–8% efficient. The remaining 92–93% of the power goes directly into producing heat. Adding to the problem is the fact that much modern program material is bass-heavy.

As understanding of heat transfer mechanisms in loudspeakers grew, designs appeared that improved heat transfer from the voice coil and gave improved thermal power handling ratings, Fig. 17-26.

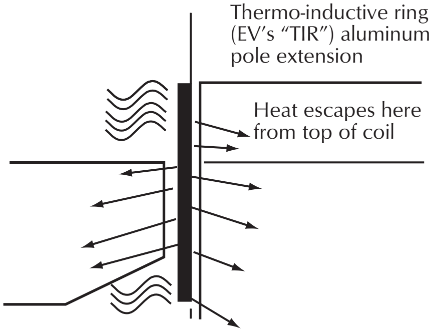
The heat transfer methods discussed here are simply methods to transfer heat away from the voice coil. If there were no heat transfer paths out of the magnetic circuit, the speaker's temperature would continue to rise without limit. In the cases of drivers on exposed horns, natural convection transfers sufficient thermal energy to prevent overheating. The thermal resistance of the direct convection transfer path is on the order of 1–2°C/W. In the case of a woofer in a fiberglass-lined enclosure, this resistance may be five times greater. Heat buildup can be substantial. This mechanism is often ignored. Several proprietary loudspeaker systems have been developed in an attempt to address this problem, but most sound reinforcement systems still provide no designed-in mechanism for transferring heat out of the enclosures.

17.7 Radiator Types

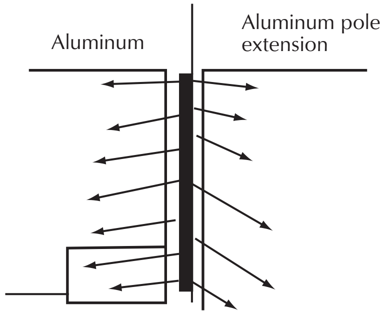
In addition to converting electrical energy to mechanical energy, a loudspeaker must include a means for converting mechanical energy (e.g., the motion of a diaphragm) into acoustic energy. For purposes of this chapter, the various means for accomplishing this final conversion are referred to as *radiators*. While there is overlap in our definition of the terms *transducer* and *radiator,* it is extremely useful in understanding loudspeakers to consider the function of acoustic radiation as a separate subject from electroacoustic conversion. In general, there are two broad types of radiators: direct radiators and horn radiators.



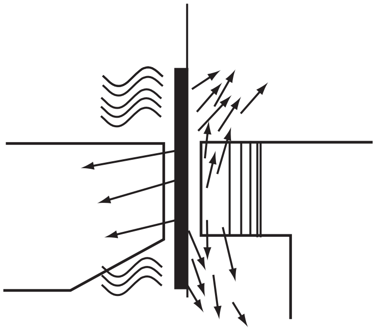
A. Straight pole heat transfer.



B. Aluminum pole extension on EV "DL" woofers.



C. Improved full-coil air conduction in EV "EVX" woofers.



D. JBL "vented gap" natural forced-convection heat transfer.

Figure 17-26. Various coil/gap geometries showing the evolution of heat transfer designs in modern woofers.

17.7.1 Direct Radiators

The simplest form of radiator is the direct radiator, in which the diaphragm is directly coupled to the air. Most hi-fi loudspeakers consist of combinations of different sizes of direct radiators. Various forms of direct radiators were described in the previous section. In this section, we will outline their acoustic attributes.

17.7.2 Cone Radiators

In a cone radiator, the diaphragm is in the shape of a truncated cone. The concave surface of the cone is usually, but not always, the one which radiates sound. The cone shape is partially dictated by expediency: it allows the magnet structure to reside at the rear of the transducer assembly, while at the same time allowing for the use of a spider and a surround to suspend the diaphragm. This dual-element suspension provides positive centering of the voice coil in the magnet gap, and it helps constrain the motion of the cone to the desired linear path.

The above notwithstanding, there are also acoustic motivations tend to favor the cone shape. At first glance, one might expect a flat piston to offer superior on-axis response and directivity to a cone. A cone shape is generally preferable, however, when the excitation will be applied near the center of the diaphragm. Due to the fact that sound propagates at a finite velocity in a solid, the motion of the outer portion of the diaphragm will follow the initial excitation by some amount of time. If the diaphragm were flat, radiation from the outer portions of its surface would arrive at an on-axis observation point at later times than radiation from the center. The cone shape reduces the distance must be traveled by sound radiated from the outer portions to on-axis listening positions. Since the velocity of sound in the cone material is typically greater than the velocity of sound in air, a cone shape having the optimum included angle will tend to synchronize on-axis radiation from the outer portions of the diaphragm with that from near the center. Given a judicious choice of angle, the useful range of response of a cone transducer can be extended to a significantly higher frequency than would otherwise have been the case.

17.7.3 Dome Radiators

Another variant on the direct radiator theme is the dome radiator. Most often, this radiator takes the form of a convex dome driven and suspended at its periphery. The material used to form the dome may be soft, as is the case with coated synthetic textiles, or hard, in the case of metals or composite materials (e.g., carbon fiber/epoxy). Dome radiators are most popular for high-frequency elements, although a number of dome-shaped midrange elements are also available. As with the cone radiator, the convex shape of the typical dome radiator has acoustic motivations. Since the excitation is at the edge, the dome's mechanical motion will propagate inward. As a result, if the shape were flat, radiation from the inner portion would arrive at an on-axis observation point later than radiation from the edge. The convex shape helps deliver a more coherent wavefront to an on-axis listening position. It is common practice to suspend a small round cover just in front of the center of the dome.

17.7.4 Ring Radiators

Yet another form of direct radiator is the ring radiator. In a typical ring radiator, a flexible ring-shaped diaphragm is rigidly captured along its inner and outer circumferences and driven along a concentric circular line between those two circles. There is, then, no distinction between the diaphragm and the suspension, as a single part fills both functions. A dome tweeter with a cover over the center of the dome functions as a ring radiator. Ring radiators can also be used to drive short horns.

The JBL 075 Bullet is an example of a ring radiator. Intended for use above 7 kHz, the diaphragm is a V-shaped ring of aluminum attached to a voice coil and former.

Fig. 17-27 shows a ceramic version of the ring, which is made by Yamaha. The phase plug is a simple slit ending in a large enough mouth to project the desired low end of the driver. The suspension is the diaphragm itself, and it is quite stiff. Ring radiators are typically operated in or above the principal resonance frequency of the diaphragm assembly.

17.7.5 Panel Radiators

Both electrostatic and *planar* electrodynamic speakers fall into this category. As with the ring radiator, there is a mixing of functionality between the diaphragm and the suspension. The acoustic advantage, at least in principle, of a panel radiator, is that the driving force is applied uniformly over a large portion of the diaphragm. For this reason, diaphragm rigidity is not an essential design element, as is the case with cone radiators. An interesting characteristic of a large panel radiator is that it will essentially project a shadow of its shape as a listening pattern; this shadow of the speaker's radiation pattern will take up a large part of a typical listening area, particularly at close listening distances. This is claimed to produce a wider "sweet spot" compared to conventional cone systems.

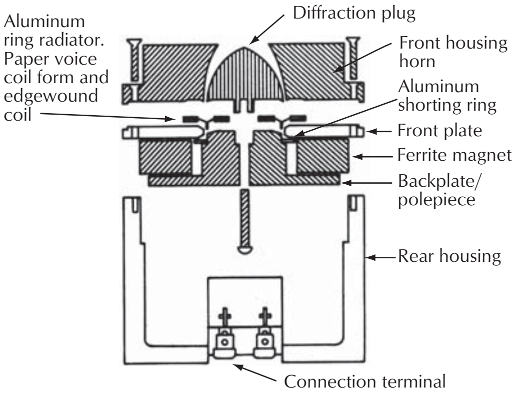


Figure 17-27. Yamaha ceramic-magnet ring radiator cross section. Courtesy Yamaha International Corp.

17.7.6 Horns

Horns are used to increase the efficiency of a transducer and to control the directivity of the sound that is radiated. Horns are characterized by a number of parameters. The earliest approach to a predictive model, and the one still employed in acoustics texts, is characterization by the rate of increase of cross-sectional area with longitudinal position in the horn. Other means of characterization are related to the shapes formed by the horn walls.

Of all possible expansion (or flare) rates, a relative few have found use in horn design and analysis. Those most commonly encountered are exponential, hyperbolic, conic, and catenary. In general, the change of cross-sectional area with position in a horn can be expressed as

$$A(x) = F(x) \tag{17-4}$$

where,

$A(x)$ is the cross-sectional area at a point x along the axis of the horn,

$F(x)$ is some function of x.

For example, in an exponential horn,

$$A(x) = A_0 e^{mx} \tag{17-5}$$

where

A_0 is the area of the horn at its throat or entry,

m is a constant called the *flare rate*.

Much more detailed information is available on the subject of the acoustic characteristics resulting from different rates of expansion from the sources cited in the Bibliography. The models are useful in analyzing the propagation of acoustic energy within a horn, but other considerations become dominant in determining the nature of the radiated sound beyond a horn's mouth.

For this reason, practical horns have come to be known more by salient details of their sidewall shapes than by their flare rates. The more common types are described below.

17.7.6.1 Radial Horns

Radial (or sectoral) horns were claimed to allow a natural radial expansion of the sound wave from the driver, while maintaining an exponential expansion rate. Typically, a radial horn has straight horizontal sides and top and bottom walls that are in the form of spherical *sectors*. The design approach employed for a radial horn involves positioning the sides at approximately the desired angle for horizontal coverage. Given the area expansion desired, the top and bottom surfaces are then derived mathematically. The most popular materials used in making radial horns are cast aluminum (now relatively uncommon), molded plastic, laminated glass fiber, and polyester resin. This type of horn was in widespread use from the 1930s until approximately the mid-1980s, by which time constant directivity types had become more popular.

Fig. 17-28 is an Altec 311-60; it has a 60° horizontal coverage and is intended for use above 300 Hz, using a 1.4 inch driver. Altec was well known for this design, with its characteristic vertical vanes at the mouth of the horn.

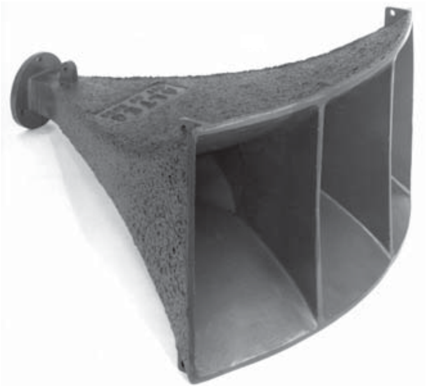


Figure 17-28. Altec Lansing 311-60 cast aluminum sectoral horn with sound-deadening material. Courtesy Altec Lansing Corp.

17.7.6.2 Multicell Horns

Multicell horns were the first horns to be employed specifically for their directivity control attributes. The design approach was straightforward — several small horns were affixed together in an array, with each horn to supply a portion of the total coverage angle. These small horns were connected to a common manifold so that a single driver could power them, Fig. 17-29. Multicell horns first came into use in the late 1930s. They were originally made of sheet metal soldered together and either filled on the outside with sand or covered with a mechanical damping material.

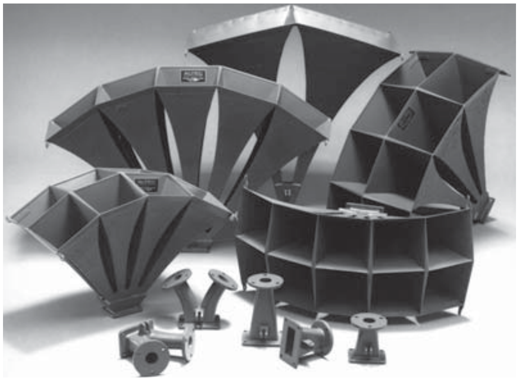


Figure 17-29. Altec Lansing 1.4-inch throat, all soldered and coated steel horn family showing throat plumbing fixtures. Courtesy Altec Lansing Corp.

17.7.6.3 Controlled Directivity Horns

The first constant directivity type of horn appeared in 1975. Developed by Electro-Voice, they employed a hyperbolic-flare throat section coupled to a conical radial bell section, as shown in Fig. 17-30. This horn shape yielded good low-frequency loading and relatively constant angular beamwidth in both vertical and horizontal directions over a wide frequency range. At the time, its design represented a major departure from previous thinking. Don Keele, the designer of the horns, presented an AES paper ("What's So Sacred about Exponential Horns"). In the paper, he disclosed several empirically developed relationships between mouth size, frequency, and maintenance of coverage angle. The concept of a waveguide as applied to an acoustic

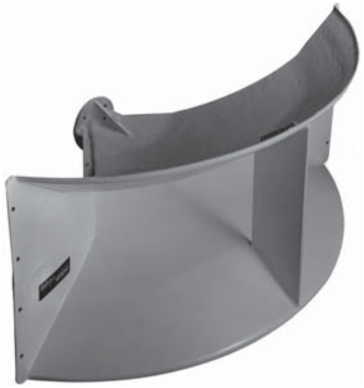


Figure 17-30. Electro-Voice HR9040 constant directivity horn. Courtesy ElectroVoice, Inc.

radiator was used as the basis for predicting and controlling the directivity of a horn.

Keele's paper and horn designs provided impetus for further empirical investigations of controlled directivity horns. Altec Lansing, at that time a competitor of Electro-Voice, introduced a family of horns using a narrow, vertical diffraction slot located at an intermediate point in the horn. With the appropriate choice of location for this slot, it is possible to make a horn with any desired combination of sidewall angles and aspect ratio (relationship between the height and width of the mouth).

This family of devices was dubbed "Manta-Ray," and a number of designs based on this thinking were introduced over the ensuing years, Fig. 17-31.

Another approach to achieving the goal of frequency-independent directivity was represented in the JBL *biradial* family of horn designs. Also developed by Don Keele, who had by then taken an engineering position with JBL, the biradial shape employs continuously varying flanges in both directions, ending in a continuous horn. The vertical diffraction slot was retained. An exponential expansion rate was part of the design, and vertical and horizontal radial bell shapes (thus the term *biradial*) were employed. Three biradial horns are shown in Fig. 17-32. They are fabricated from cast aluminum (throat section) and molded fiberglass (bell section) and fit 2 inch exit drivers.

17.7.6.4 Voice Warning Horns

Fig. 17-33 shows another variety of controlled directivity horn, designed by Bruce Howze of Community Professional Loudspeakers, and originally built for Whelen Engineering. The horn used a slightly different directivity control philosophy: a controlled horizontal pattern (45° and a narrow vertical pattern), due to the 70 in (1.78 m) vertical dimension. The horn uses 16 siren drivers and the system, with a 1600 watt input, generates 127 dB at 100 m (328 ft). It is used in a similar manner to a searchlight: aim and shoot via a remote rotor.

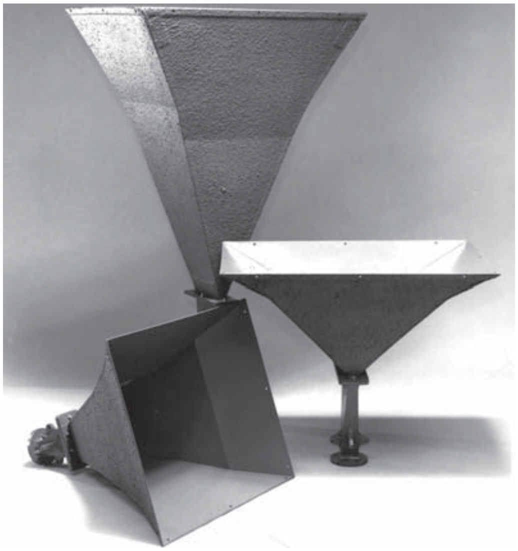


Figure 17-31. Altec Lansing Manta-Ray horn family, cast aluminum throat and soldered, coated bell construction. Courtesy Altec Lansing Corp.

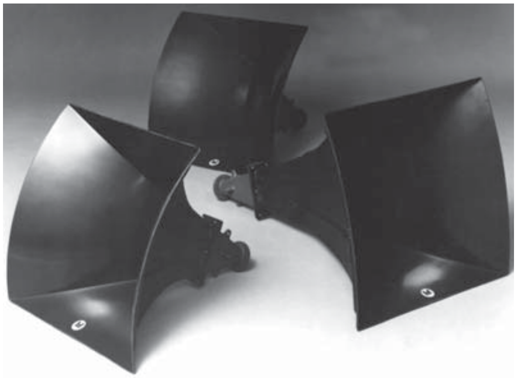


Figure 17-32. JBL biradial horn family cast aluminum throat and fiberglass bell construction. Courtesy JBL/UREI.

17.7.6.5 Asymmetric Directivity Horns

As the ability to determine the optimum loudspeaker directivity requirements for specific applications was refined, it became apparent that the required directivity was usually not symmetrical about a horizontal plane through the axis of the horn. From the point of view of

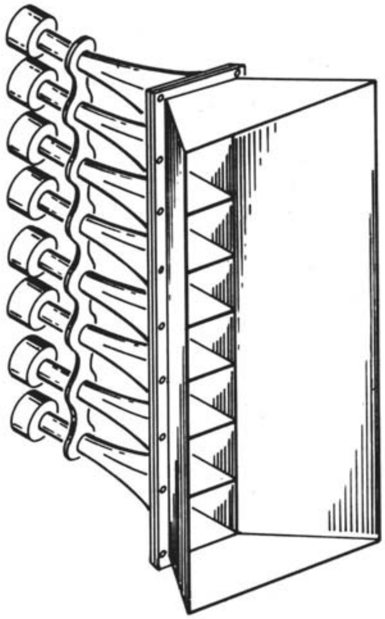


Figure 17-33. Whelen Engineering horizontal diffraction horn with multiple drivers. Courtesy Whelen Engineering.

the loudspeaker, it is most common for the required horizontal coverage to be relatively narrow at the greatest distance from the source and to become successively wider at closer distances. In addition, it is desirable for the vertical angle of greatest intensity to be as large as possible — i.e., for greater energy to be directed to the seats at the greatest distance from the loudspeaker in order to produce similar SPL values throughout the audience. One early attempt to address this requirement was the JBL 4660, shown in Fig. 17-34.

Another design, developed for a specific application, is the IMAX® PPS (Proportional Point Source) loudspeaker, developed by the author. Fig. 17-35 is the high-frequency horn used in this loudspeaker, and Figure 17-36 is its 4 kHz isobar.

Dave Gunness, chief engineer at Electro-Voice at the time, developed a family of asymmetric directivity horns in the late 1980s and early 1990s. These were known as *Vari-Intense* devices.

Optimized (asymmetric) directivity is an attractive engineering goal, but there are a number of obstacles to its widespread acceptance:

1. Computer-based sound system prediction software is required in order to visualize its effectiveness and optimize aiming and device placement,
2. Exactly what constitutes ideal directivity is a strong function of the space in which the device is to be used. The ideal directivity will vary, for example, for different loudspeaker elevations within the same space.
3. With the exception of the proprietary IMAX loudspeaker, there are currently only high-frequency devices available with this type of directivity. Achieving uniform sound pressure levels throughout the seating, but at high frequencies only, is of limited value.

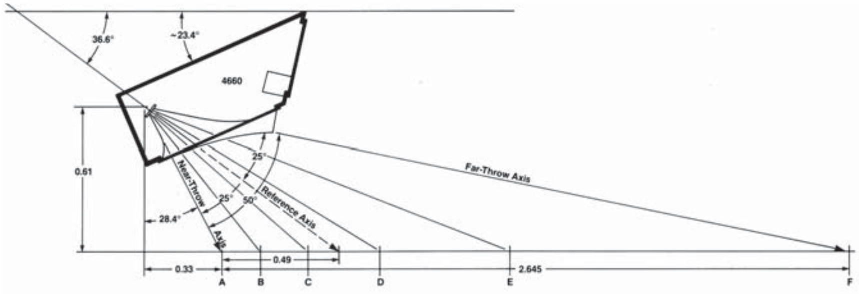
17.7.6.6 Acoustic Lenses

Although an acoustic lens is not generally regarded as a directivity control device, it can function as a directivity alteration device. While acoustic lenses are used to widen a pattern, they can also be used to narrow a horn's directivity. An acoustic lens is usually formed with parallel plates of strategically chosen shapes placed at an angle to the direction of sound propagation. Differing path lengths through different portions of the lens create arrival time relationships for the associated components of the wave that generate specific directivity characteristics.

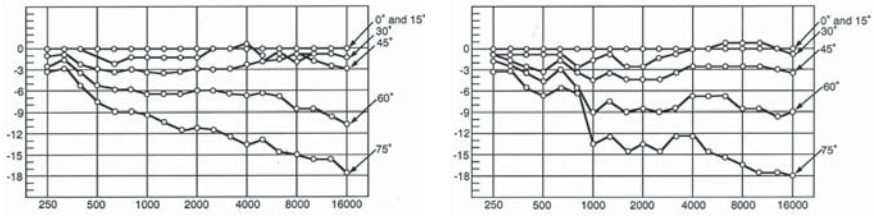
The slant-plate lens assembly, shown mounted on a JBL studio monitor in Fig. 17-37, is one notable implementation of an acoustic lens. Note that the device has concave openings in the plate array. As the wave leaves the horn and progresses through the lens plate array, the center of the wave reaches the air on the outside first, due to the shorter path through the lens. The outer portions of the wave travel through longer paths within the lens and are therefore delayed in time relative to the portions that came from the center. The net effect is to produce arrivals that are better synchronized—and therefore stronger—at positions that are off axis in the horizontal, thus widening the polar pattern in that direction. The vertical pattern would ideally be unaltered. Lenses have the undesirable property of causing relatively strong reflections back into the horn.

17.7.6.7 Folded Horns

One of the practical drawbacks of horns, particularly those intended for use at low frequencies, is their physical size. Folded horns were developed in response to this problem and have been in use in various forms for more than half a century. A folded horn is produced by truncating the shape at point, providing a reflecting surface to change the direction of the outgoing wave, and continuing the horn's expansion in another direction,

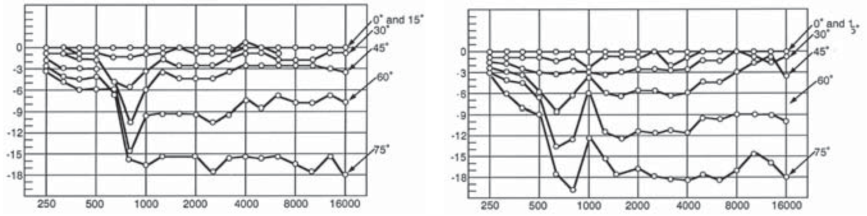


A. Standard orientation.



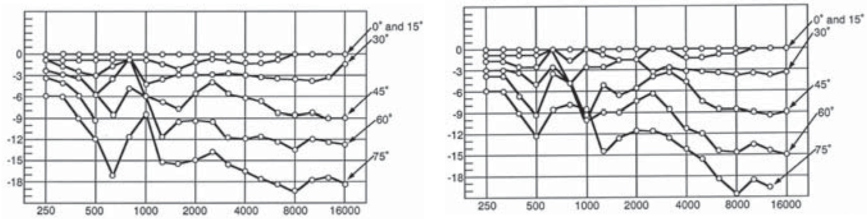
B. Off-axis response at A.

C. Off-axis response at B.



D. Off-axis response at C.

E. Off-axis response at D.



F. Off-axis response at E.

G. Off-axis response at F.

Figure 17-34. JBL 4660 Asymmetric-directivity horn. Courtesy JBL.

usually opposite the prior one. Successive horn sections are typically positioned outside their predecessors. As many reversals are generated as are necessary to create the desired path length and mouth size. There are, as with other horn types, many variations on folded horns.

Fig. 17-38 shows a University Sound GH directional trumpet cross section and how the area expands by making two 180° turns. The design was introduced in the 1940s. Another University folded public address horn design is the cast zinc Cobreflex, shown disassembled in Fig. 17-39. This horn expands into a double mouth.

One of the most recognizable low-frequency horns is the Klipschorn, shown in Fig. 17-40. It was named for its inventor, Paul Klipsch, who was one of the pioneers in horn loudspeaker design. The Klipschorn uses a single 15 inch loudspeaker in a relatively compact package. It is designed for placement in a corner of the room, with the room's walls forming an extension of the horn shape.

The Cerwin-Vega E horn is another form of folded bass horn. An 18 inch low-frequency driver sits between the upper and lower mouths and faces to the

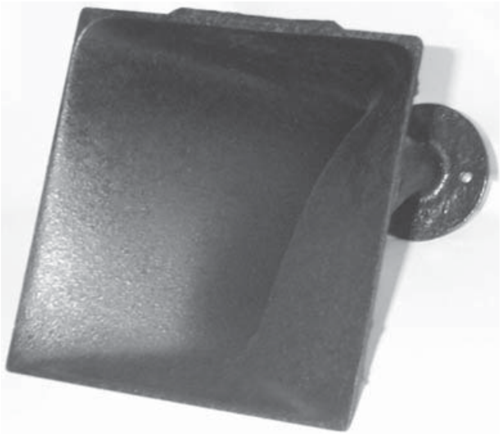


Figure 17-35. IMAX Proportional Point Source high frequency horn. Courtesy James E. Mitchell & Associates.

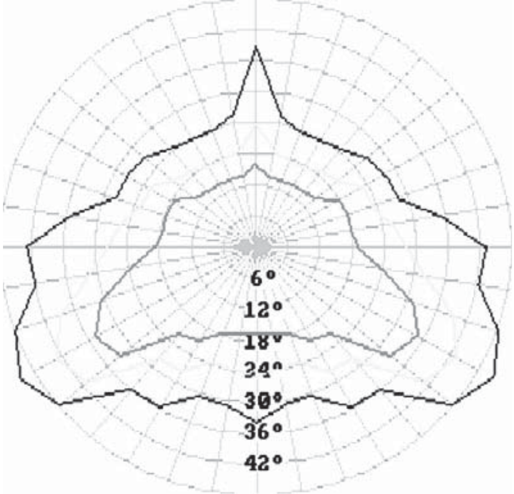


Figure 17-36. 4 kHz isobar of the loudspeaker in Fig. 17-35. Courtesy James E. Mitchell & Associates.

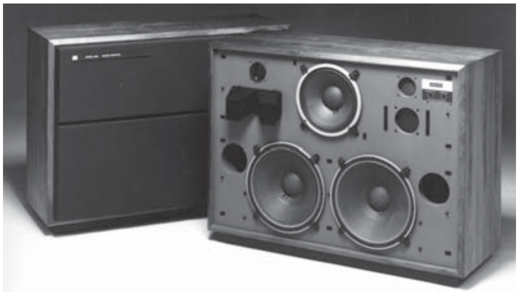


Figure 17-37. JBL studio monitor employing slant-plate-type acoustic lens on a high-frequency component. Courtesy JBL.

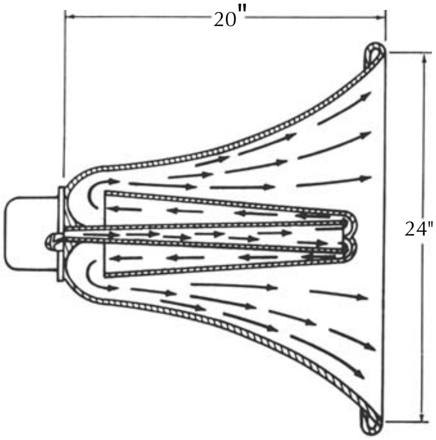


Figure 17-38. Cross-sectional drawing of an exponential folded horn.

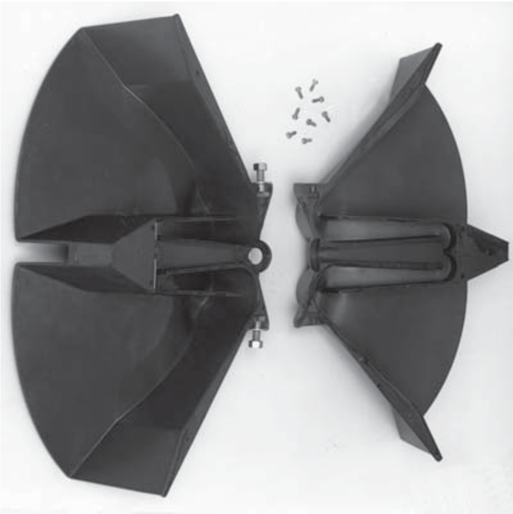


Figure 17-39. University Cobraflex horn disassembled. Courtesy Altec Lansing Corp.

rear in a compression chamber. The E horn makes one 180° fold that opens to the double mouths. The horn is intended for use with additional mouth extensions and in multiples for low-frequency coupling.

The W horn is another folded bass horn design. The best-known example is the RCA theater horn, which uses two 15 inch drivers. The W uses forward-facing drivers, and the horn flare is designed as a W-shaped double fold, expanding to twin mouths.

The Altec 31A uses a single 90° fold. This allows for a short front-to-back dimension. Additionally, the driver is mounted facing downward, making it rain and dust resistant. This horn uses a 120º mouth and is used above

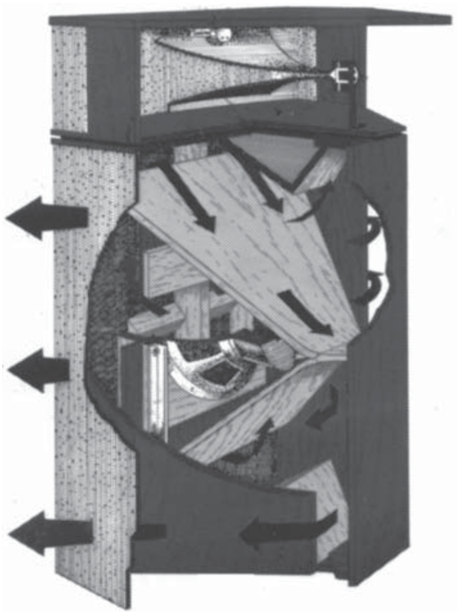


Figure 17-40. Klipschorn folded corner bass horn—rear cutaway view with high-frequency components. Courtesy Klipsch & Associates.

500 Hz for a wide variety of applications, including voice-only systems.

The obvious advantage of a folded horn is the reduced package size for a given horn length. This advantage is offset by the fact that, for each reversal fold in the horn's shape, a reflection is generated inward, opposite the desired direction of wave propagation. These reverse waves are reflected again in a forward (outgoing) direction when they reach the horn's driver area, generating late signal arrivals that cause significant deviations from ideal in the horn's response. For this reason, folded horns generally find use in applications that are relatively undemanding of fidelity.

17.7.6.8 Special Considerations for Low-Frequency Horns

Over the years, a number of horns have been developed specifically to radiate low frequencies. In the past, the primary motivation for the use of a horn to reproduce low frequencies was improved efficiency as compared to a direct radiator. The current availability of power amplifiers with extremely high output capacities and woofers that are capable of utilizing that power has rendered the issue of efficiency less important than that of size. As a result, there are fewer low-frequency horns on the market today than in the past.

The most common difference between bass horns and those intended for mid- and high-frequency use, aside from the bass horns' larger size, is that low-frequency horns typically do not employ compression drivers. Instead, a cone transducer is mounted directly in the throat of the horn.

A potential issue in low-frequency horn design and operation is the transitional behavior of the horn. It is common practice to use a bass horn/driver combination to a sufficiently low frequency that the horn is too small to provide substantial acoustic loading in the lower portion of the bandwidth of use. In this frequency range, the driver must operate as a direct radiator, with correspondingly lower sensitivity. Although it has been asserted that this discrepancy, which can exceed 10 dB, may be overcome through the use of ports, in actuality the only means of leveling the device's response between the two regimes of operation is with equalization of the input signal. Given proper equalization, a bass horn may be used in this fashion with excellent results.

17.8 Loudspeaker Systems

Most practical loudspeakers are systems comprising multiple transducer/radiator subsystems, each of which radiates a portion of the audio-frequency spectrum. This area of loudspeaker design has a major impact on a loudspeaker's ultimate performance, yet this portion of the design process is frequently shortchanged. In this section we will discuss some considerations for loudspeaker system design and performance and provides some illustrative examples.

The desirability of dividing the audible frequency range into multiple bands is taken for granted in most loudspeaker applications. The most compelling reasons for dividing the spectrum among multiple components are:

1. By itself, the bandwidth of a practical transducer/radiator is inadequate to meet the bandwidth requirements for a complete loudspeaker.
2. The directivity of a single transducer/radiator will not be sufficiently consistent with frequency to meet reasonable goals for the directivity of a full range loudspeaker.
3. The maximum available acoustic output of a single transducer is inadequate. Sharing the output demand among a number of band-specific components enables a loudspeaker to produce greater total acoustic power.

In designing a loudspeaker system, one should, at the very least, have a working knowledge of the disciplines involved in the design of the component parts. The system designer's challenge is to make a collection of individual components function as a cohesive whole while meeting the cost, size, and aesthetic requirements of the loudspeaker's intended applications. The design of a successful loudspeaker system involves much more than simply selecting a group of components and building a box to house them.

It is axiomatic that, in addition to the required technical expertise, a loudspeaker designer should have the capability of subjectively evaluating a loudspeaker's performance—critical listening—and that the final determinant of a loudspeaker's success will almost always be subjective acceptance. It is equally true that there are always objectively observable phenomena that correlate with subjective preferences. The difficulty in reconciling the two is a direct result of the very large number of objective elements that must be accounted for in order to fully characterize the performance of a loudspeaker. This subject is covered in greater depth in the "Loudspeaker Characterization" section of this chapter.

Loudspeaker systems are often categorized by the number of spectral divisions made in the system, as in two-way or three-way systems. Generally speaking, a loudspeaker system consists of two or more transducer/radiator combinations, a crossover network, and an enclosure that houses everything. In addition to providing a convenient package for the components, the enclosure serves structural, acoustic, and aesthetic purposes. The sections on acoustic boundaries and electroacoustic models provide information about some of the acoustic effects of enclosure design.

17.8.1 Configuration Choices

A number of decisions about a loudspeaker's configuration are typically made early in the design process. These include:

1. The number of spectral bands, or divisions.
2. The type of radiator to be used for each band.
3. The location and orientation of the individual components within the system housing.

In determining the number of frequency bands to be used in a loudspeaker, several conflicting demands must be reconciled. Choosing a greater number of divisions creates the possibility of greater broadband acoustic output and more optimal radiator configurations for each band. On the other hand, each added band adds to the size, complexity, cost, and more often than not, to non ideal aspects of the acoustic behavior of the finished design.

The type of radiator chosen for each band is often a matter of custom or convention rather than of engineering. Where possible, it is generally desirable to match efficiencies and directivities of adjacent bands over a range of frequencies centered about their crossover point. This is most readily accomplished when similar types of radiators are used for both of the bands in question.

The location and orientation of individual components is an area worthy of careful attention. It is common practice to place all of the transducers on a flat panel (a baffle), displaced from each other in vertical and/or horizontal directions. In the case of loudspeakers designed for stereo reproduction, it is also common practice to make pairs of speakers in a mirror-image layout.

The aforementioned common practices have developed over many years, with the primary motivation being cost and ease of manufacture. Another approach to loudspeaker system design is the coaxial layout. First employed in the earlier part of the 20th century, this practice involves locating two or more bands of a loudspeaker along a common axis. While the coaxial approach is typically more difficult to implement, it has some advantages over more conventional layouts.

17.8.2 Types of Loudspeaker Systems

The simplest form of loudspeaker employs a single full-range transducer to reproduce all frequencies. The most common applications for this type of device are limited-bandwidth (e.g., speech) systems and inexpensive music reproduction systems.

For residential music systems, one of the more common configurations is a two-way system utilizing a small (typically 6 or 8 inch) woofer and a dome tweeter. Fig. 17-41 is one such system. More elaborate (and costly) systems are also employed for residential use, some employing line arrays (see Section 17.8.4) of transducers for one or more of their bands. As is the case with professional loudspeakers, visual aesthetics can play as important a role as performance in setting design requirements.

One of the more common types of loudspeakers for general sound reinforcement use is a two-way system consisting of a direct radiator woofer and a horn/compression driver high-frequency subsystem, with the components being located one above the other on the front face of the enclosure. The typical package for this type of loudspeaker is a trapezoidal enclosure.

GENELEC
GENELEC

Figure 17-41. Two-way monitor speaker. Courtesy Genelec.

The trapezoid designation describes the plan view of the enclosure, and the shape allows multiple loudspeakers to be arrayed in the shape of an arc segment, with the included angle between adjacent loudspeakers being equal to twice the sidewall angle. A large number of manufacturers offer loudspeakers that fit this description. Typical sidewall angles range from 12° to 15°, while the horizontal coverage angle (the angle at which the output has fallen to 6 dB below the level on axis) of the high-frequency horns used in such devices is typically either 60° or 90°, and the coverage of the woofer is entirely uncontrolled.

17.8.3 Performance Issues in Multiway Systems

Before trying one's hand at designing a multiway loudspeaker, it is a good idea to develop a familiarity with the complete audio signal chain and to understand the implications each decision will have on the acoustic signal that will reach the listener's ear. Fig. 17-42 is a functional diagram of the electrical and acoustic signal paths from the source (electrical input signal) to the observation (listening) point.

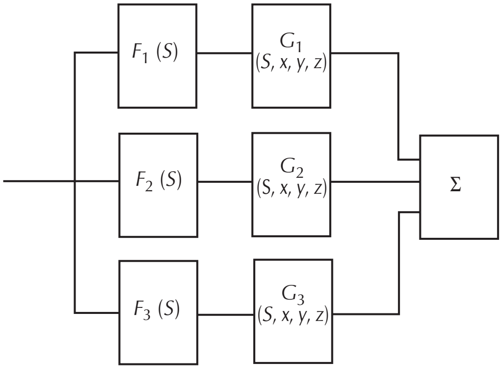


Figure 17-42. Functional diagram of multiway loudspeaker.

In the above representation,

$$F_T(S, x, y, z) = \Sigma_N[F_n(S)G_n(S, x, y, z)] \tag{17-6}$$

where,

F_T is the total electroacoustic transfer function,

F_n is the (electrical) transfer function of the n^{th} crossover filter,

S is the Laplace complex frequency variable,

G_n is electroacoustic transfer function of the nth radiator in the system,

x, y, and z are Cartesian spatial coordinates.

The concept is general and will accommodate an arbitrary number of spectral divisions.

It should be noted that, although this diagram depicts a loudspeaker with a passive crossover, it might also be used to represent an active system simply by including gain in the transfer functions of the crossover blocks. An active loudspeaker simplifies the task of crossover design. Since the power amplifier serves as a buffer between the crossover and the transducers, the frequency-dependent impedance behavior of the transducers becomes a very small factor in the design process. Note that the crossover filters are in cascade with the transducers, while the acoustic outputs of the devices are summed acoustically at the listening position. The nature of such a multipath system is complex, and detailed prediction of its response at every likely listening position is a nontrivial task.

Note also that the transfer functions G_n are functions of the spatial coordinates x, y, and z as well as of the complex frequency variable S. This spatial dependency includes the effects of source directivity, propagation delay, and the inverse square law. If the system is not coaxial (and sometimes even if it is), then the lengths of the paths from each transducer to a given listening position will not generally be the same. The most common practice is to choose an axis along which one will attempt to equalize these acoustic path lengths and then to optimize the speaker's behavior on this axis. In the case of a two-way loudspeaker with the transducers displaced along a line in the plane of the baffle, it is possible to make path lengths equal at every point in a plane. Once more than two spectral divisions are present, even this limited goal is no longer possible. It may well be the case that, in a three- or four-way system, there is no point for which all acoustic path lengths from the transducers to a listener's ears will be equal.

The effect of unequal acoustic path lengths is that signals from different radiators will reach the listener at different times, even though they originated simultaneously. This timing discrepancy is not generally sufficient to be recognized by a listener as comprising distinct multiple events, but it is enough to have audible and undesirable effects on a loudspeaker's amplitude response, as well as on its ability to reproduce transient signals. Even when an axis or plane exists in which signal synchronization has been achieved, positions off the axis or outside of the plane will not receive the benefits of such synchronization.

The subject of complex addition of time-varying signals is beyond the scope of this chapter, but it is dealt with in many introductory circuit analysis texts. Additional effects of signal synchronization, and the lack thereof, on loudspeaker response are illustrated in the section on crossovers.

There are a number of ways in which the problems caused by noncoincident transducer locations may be addressed by a loudspeaker designer. One common approach is simply to assert that the response anomalies caused by this configuration are not audibly significant and to accept (or avoid acknowledging) their presence. Another is to employ crossover filters with very steep slopes so as to minimize the frequency range over which anomalies due to path length differences will be present. As will be shown in the crossover section, the latter technique has its own set of drawbacks and may in some cases create more serious problems than it solves.

One means of addressing the synchronization issue is with a coaxial loudspeaker. This type of loudspeaker is most often two-way, although it is possible to design a three- or four-way coaxial system. There are benefits in making the midrange and high-frequency components of a three-way system coaxial, while leaving the low-frequency portion displaced in the more conventional manner.

A coaxial loudspeaker will always possess symmetrical response behavior, Fig. 17-43. That is, the response at a given angle from its axis will be mirrored at the same angle in the opposite direction. Additionally, if the acoustic path lengths from transducer to listener are equal on the system's axis, it is possible to preserve this synchronization at all listening positions with a coaxial design. Even though it is possible to achieve signal synchronization over a wide angular range with a coaxial loudspeaker, this possibility is not always realized in practice. When a coaxial loudspeaker fails to achieve coincident performance (i.e., it fails to behave as a single full range radiator), its sole distinction as compared to more conventional configurations is that frequency-dependent anomalies related to crossover interactions will be symmetrically located about the loudspeaker's axis.

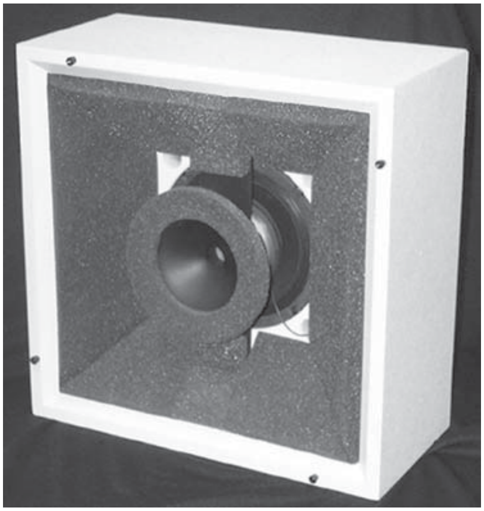
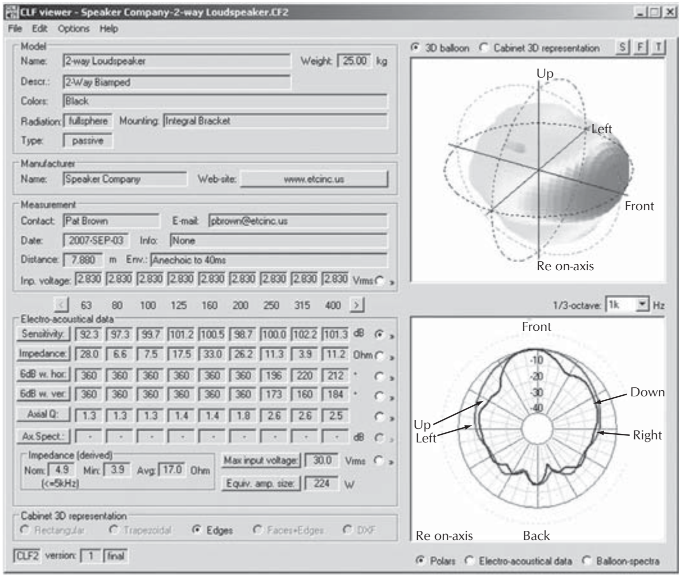


Figure 17-43. Small two-way coaxial loudspeaker. Courtesy Frazier Loudspeakers.

Fig. 17-44 illustrates some of the effects caused by displaced transducers. The frequency selected for display is the closest ⅓-octave band center to the crossover frequency of the speaker. Figure 17-45 shows the improved polar behavior that can be produced with a coaxial loudspeaker, again with the ⅓-octave band center chosen so as to most closely match the crossover frequency. Note that the polar pattern of this speaker is still not perfectly symmetrical, even though the transducers are coaxial. This is due to asymmetric placement of the coaxial woofer/tweeter assembly within the enclosure. The effects of enclosure design on loudspeaker performance will be covered in more detail in the section on acoustic boundaries.

17.8.4 Line Arrays

Another type of loudspeaker system is a line array. Although line arrays have much in common with other types of loudspeaker systems, they have some attributes that are unique enough to justify their separate treatment. A line array may form a complete full-range loudspeaker or one or more bands thereof. In a line array, individual radiators are arranged in a straight line or an arc segment. It is also possible for a number of complete loudspeaker systems to be configured as a line array. It is this configuration that has come into fashion in recent years. In the simplest form of line array, each of the elements—usually a small cone transducer—is supplied an identical full-range signal. This type of array, also called a *sound column*, was popular in this



CLF viewer - Speaker Company-2-way Loudspeaker.CF2

Model
Name: 2-way Loudspeaker Weight: 25.00 kg
Descr.: 2-Way Biamped
Colors: Black
Radiation: fullsphere Mounting: Integral Bracket
Type: passive

Manufacturer
Name: Speaker Company Web-site: www.etcinc.us

Measurement
Contact: Pat Brown E-mail: pbrown@etcinc.us
Date: 2007-SEP-03 Info: None
Distance: 7.880 m Env.: Anechoic to 40ms

	63	80	100	125	160	200	250	315	400	
Inp. voltage:	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	Vrms ○
Sensitivity:	92.3	97.3	99.7	101.2	100.5	98.7	100.0	102.2	101.3	dB ●
Impedance:	28.0	6.6	7.5	17.5	33.0	26.2	11.3	3.9	11.2	Ohm ○
6dB w. hor:	360	360	360	360	360	360	196	220	212	° ○
6dB w. ver:	360	360	360	360	360	360	173	160	184	° ○
Axial Q:	1.3	1.3	1.3	1.4	1.4	1.8	2.6	2.6	2.5	○
Ax.Spect.:	-	-	-	-	-	-	-	-	-	dB ○

Impedance (derived)
Nom: 4.9 (<=5kHz) Min: 3.9 Avg: 17.0 Ohm
Max input voltage: 30.0 Vrms
Equiv. amp. size: 224 W

Cabinet 3D representation
○ Rectangular ○ Trapezoidal ● Edges ○ Faces+Edges ○ DXF

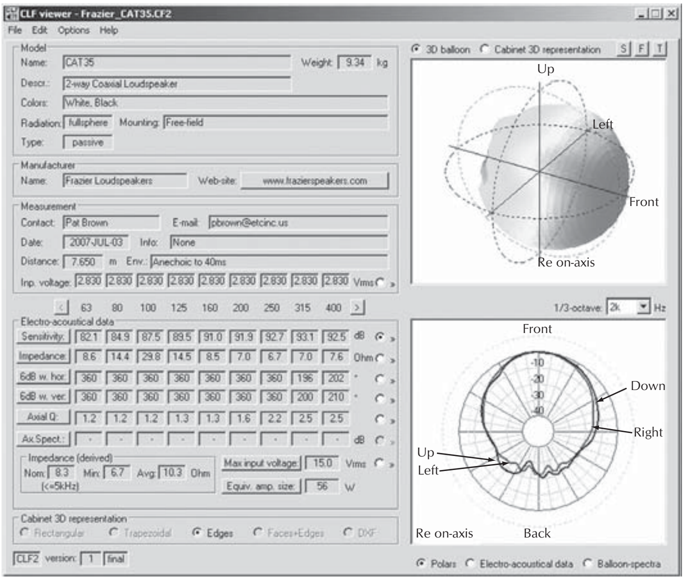
CLF2 version: 1 final

● 3D balloon ○ Cabinet 3D representation

1/3-octave: 1k Hz

● Polars ○ Electro-acoustical data ○ Balloon-spectra

Figure 17-44. Generic two-way loudspeaker directivity balloon and polar pattern in the crossover region.



CLF viewer - Frazier_CAT35.CF2

Model
Name: CAT35 Weight: 9.34 kg
Descr.: 2-way Coaxial Loudspeaker
Colors: White, Black
Radiation: fullsphere Mounting: Free-field
Type: passive

Manufacturer
Name: Frazier Loudspeakers Web-site: www.frazierspeakers.com

Measurement
Contact: Pat Brown E-mail: pbrown@etcinc.us
Date: 2007-JUL-03 Info: None
Distance: 7.650 m Env.: Anechoic to 40ms

	63	80	100	125	160	200	250	315	400	
Inp. voltage:	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	Vrms ○
Sensitivity:	82.1	84.9	87.5	89.5	91.0	91.9	92.7	93.1	92.5	dB ●
Impedance:	8.6	14.4	29.8	14.5	8.5	7.0	6.7	7.0	7.6	Ohm ○
6dB w. hor:	360	360	360	360	360	360	360	196	202	° ○
6dB w. ver:	360	360	360	360	360	360	360	200	210	° ○
Axial Q:	1.2	1.2	1.2	1.3	1.3	1.6	2.2	2.5	2.5	○
Ax.Spect.:	-	-	-	-	-	-	-	-	-	dB ○

Impedance (derived)
Nom: 8.3 (<=5kHz) Min: 6.7 Avg: 10.3 Ohm
Max input voltage: 15.0 Vrms
Equiv. amp. size: 56 W

Cabinet 3D representation
○ Rectangular ○ Trapezoidal ● Edges ○ Faces+Edges ○ DXF

CLF2 version: 1 final

● 3D balloon ○ Cabinet 3D representation

1/3-octave: 2k Hz

● Polars ○ Electro-acoustical data ○ Balloon-spectra

Figure 17-45. Coaxial two-way loudspeaker directivity balloon and polar pattern in the crossover region.

country through much of the 1970s and is still in common use in installed sound systems.

Recent developments in DSP technology, combined with the constant pressure on the touring concert reinforcement industry to minimize weight, blockage of audience sight lines by speakers, and truck space, have resulted in a resurgence of interest in line arrays. As attractive as some of their perceived performance characteristics may be, they have inherent limitations. First, the directivity attributes associated with line arrays are present in the vertical plane (along the length of the array) only. The horizontal directivity is only as good as the horizontal performance of the individual devices used to form the array. Secondly, line arrays invariably comprise discrete elements, as opposed to a continuous line source. This periodicity exacerbates problems with nulls and lobes, and it causes the off-axis impulse response of a line array to contain multiple discrete arrivals.

It is often incorrectly asserted that a line array behaves, or can behave, as a line source. A line source is largely a theoretical construct. It consists of a long, narrow radiator that radiates sound with perfect uniformity at every point on its surface. This assumption of perfect uniformity, while impossible to achieve in practice, simplifies the mathematics required to model the behavior of a line source. When used for illustrative purposes in texts, line sources may additionally be assumed to have infinite length, making possible even further simplification of the mathematical model. The same model has been employed in texts on electromagnetic theory, for the same reasons.

The two assumptions—continuous radiation and infinite length—lead to two interesting results. First, due to symmetry, the frequency response of an infinitely long, continuous line source is not a function of observation position along the line. For example, if the line is assumed to be coincident with the Z-axis in a cylindrical polar coordinate system, then its response will not vary with changes in the Z-coordinate of an observation position (i.e., for movement in a direction that is parallel to the line). Second, due to the infinite length of the source, the wavefront (a collection of isophase points) will form a cylindrical, rather than a spherical, shape. For this reason, the intensity of radiation in the outward direction falls off as the inverse of the first, rather than the second, power of the distance from the line.

As interesting and attractive as the two above results may be, they are not achievable in any physically realizable array. The effects of radiation that is neither continuous nor uniform, and of finite array length, cannot be neglected in discussing the behavior of real-world systems. Unfortunately, these issues have been glossed over or completely ignored in the information that is provided regarding the performance of commercially available line array products.

Full-range line arrays characteristically have relatively narrow vertical radiation patterns. The details of these radiation patterns vary widely with frequency and typically contain undesirable off-axis nulls (deep response notches) and lobes (response peaks). The same phenomena that produce off-axis response variations in a noncoaxial, multiway loudspeaker—interference caused by variations in the relative distances between multiple sources and the listener—create this directivity. At high frequencies, the angular separation between the first two nulls—and therefore the useful coverage angle—may be on the order of 5° or less.

A number of remedies to the problem associated with line arrays have been implemented over the past 50 years. There are two primary areas in which the line array intrinsically poses challenges to the designer: total array length and individual device spacing. Both must be addressed in order to produce a well-behaved system.

One means to address the issue of total array length is to implement a tapered array. In this type of array, only the innermost elements carry the highest frequencies. The signals applied to the more outwardly placed elements in the array are low-pass filtered at successively lower frequencies. The goal of this approach is to make the effective length of the line array become shorter at higher frequencies. An alternative way of stating this goal is that one desires the ratio between the effective length of the array and the wavelength of sound to be invariant. With the ability via DSP processing to create filters of essentially arbitrary amplitude and phase response, it has become relatively straightforward to create tapered arrays. Additionally, the availability of frequency-independent delay makes lobe steering possible.

The matter of device spacing poses another set of challenges. The smaller the spacing can be made relative to wavelength, the better a line array can approximate the behavior of a continuous radiator. When device spacing becomes large relative to a wavelength—roughly in the range of a full wavelength—the off-axis response of the array will contain many lobes and nulls. It is likely that one or more of these off-axis lobes will approach the level of the on-axis radiation. When one considers the small wavelengths of the higher audible frequencies—the wavelength of 10 kHz is 34.4 mm (1.35 inch)—the challenge of achieving optimal device spacing for higher frequencies becomes apparent. The continued reduction in size of motor assemblies through the use of high-powered magnetic materials has been helpful in addressing this issue.

17.8.5 Crossovers

Multiway loudspeakers incorporate a crossover network. A crossover network is a collection of electrical filters, each of which allows a specific portion of the frequency spectrum to pass through it. The filtered signal is then applied to one of the bands in the loudspeaker. The types of electrical filters used to execute the crossover function are low pass, high pass, and bandpass.

The simplest crossover network consists of a low pass and a high-pass filter for use in a two-way loudspeaker. Choices that must be made regarding the filters in this crossover are:

1. Crossover frequency: below this frequency, output from the low-frequency section (woofer) is dominant, and above it the high-frequency section (tweeter) dominates.
2. Filter slopes: analog filters have characteristic stopband, or rolloff, slopes, which are integer multiples of 6 dB/octave (or equivalently 10 dB/decade). The simplest type of filter is the first order, or 6 dB/octave filter. In passive loudspeakers, the highest order filters in common use are third-order (18 dB/octave), whereas fourth-order (24 dB/octave) Linkwitz-Reilly filters are popular in active crossover implementations.

17.8.5.1 Effect on Maximum Output

The choice of filter slopes used in a crossover has a number of implications for the performance of the loudspeaker system. Generally speaking, crossover filter characteristics will affect a loudspeaker's maximum output capacity, amplitude and phase response, and directivity.

Since all transducers have a maximum excursion beyond which their output is no longer linear (or permanent damage occurs), and since the required excursion for a given acoustic output level increases with decreasing frequency, the characteristics of the high-pass filter(s) in a crossover have a direct bearing on a loudspeaker's maximum available acoustic output: in general, selecting a higher cutoff frequency will reduce the excursion required of the high-frequency transducer(s), as will employing steeper filter slopes. For a given high-frequency transducer, increasing the crossover frequency reduces the displacement required of that transducer. The demand made of the woofer as a result of the increase is strictly thermal, since the lower end of its band of use is not affected by such a change. This benefit has to be balanced against the possible inability of the woofer to effectively radiate higher frequencies over a large angle.

In addition to excursion limiting, the bandwidth of the signal applied to a given transducer determines the thermal load the transducer will see in operation. For this reason, dividing the spectrum into a greater number of bands—thereby reducing the total power that is applied to any single band—can also increase the available acoustic output of a loudspeaker. One must consider, however, that very seldom will the signal applied to a loudspeaker contain a constant broadband spectrum. At times, much of the energy applied to a loudspeaker may be confined to a relatively narrow range of frequencies. In such cases, the advantage of having a greater number of loudspeaker bands is substantially reduced.

17.8.5.2 Effect on Loudspeaker Response

The choice of filter slopes and alignments has major implications for the response of a multiway loudspeaker. Even though these effects have been examined and published for decades, they are often either misunderstood or simply ignored by loudspeaker designers.

It is a good idea to state as simply as possible the ideal functional requirements that should be met by a crossover network: a crossover should enable the acoustic sum of the individual transducers' outputs to be an accurate replica of the system input signal.

The response of a loudspeaker, for purposes of this chapter, is defined as its pressure response at a particular point in space. Even though the above criterion is simple to state, there are many design constraints that lead to tradeoffs in a loudspeaker's accuracy. For example, prevention of damage to transducers is often an overriding consideration in the design of a crossover. This may motivate the designer to consider steeper filter slopes. In some loudspeaker configurations, off-axis response anomalies are intrinsic to the design. The designer may wish to make off-axis anomalies in amplitude response as geometrically symmetrical and as narrowband as possible. The Linkwitz-Reilly filter family is sometimes employed in pursuit of these goals.

The simplest crossover is a first-order filter pair. In a two-way loudspeaker, the first-order transfer functions for low-pass and high-pass functions are:

$$F_1 = \frac{\omega_0}{S+\omega_0} \tag{17-7}$$

and

$$F_h = \frac{S}{S+\omega_0} \tag{17-8}$$

where,
F_l is the low-pass transfer function,
$\omega_0 = 2\pi f_0$ is the angular cutoff (–3dB) frequency,
S is the Laplace complex frequency variable,
F_h is the high-pass transfer function.

If we add the two electrical transfer functions, we get

$$\begin{aligned} T_t &= \frac{S+\omega_0}{S+\omega_0} \\ &= 1 \end{aligned} \tag{17-9}$$

The two transfer functions add up to a constant, independent of frequency. This is a desirable result, since the outputs of the radiators in a multiway loudspeaker are ultimately recombined by (acoustic) addition. The transfer function of our electrical sum implies that, in a two-way loudspeaker with ideal, perfectly coincident transducers and a first-order crossover, the system transfer function would not depend on frequency. We could, with some additional effort, engage in the same exercise with higher-order transfer functions. If we did so, we would find that, of all symmetrical (identical low-pass and high-pass slope and alignment class) filters, only the first-order pair does not introduce phase or amplitude error or both to the loudspeaker's transfer function. The interested reader will find detailed mathematical analyses of the various crossover topologies in the references cited at the end of this chapter.

One way of examining the effects of crossover filters on loudspeaker response is to use circuit simulations to model various aspects of the system's behavior. This method has the advantage of presenting a simple graphic representation of the values being modeled without requiring extensive mathematical skills for comprehension.

17.8.5.3 Two-Way Crossovers

For simplicity, we will examine several aspects of crossover performance in two-way systems. Then we will point out some of the elements that must be altered when three- or four-way systems are contemplated.

The chart in Fig. 17-46 is the impulse response of a first-order crossover, including the input signal, low-pass, high-pass, and summed signals. For simplicity, the crossover frequency has been set at 1 kHz. The choice of crossover frequency causes no loss of generality.

Note that, although low-pass filter has obvious delay and the high-pass filter overshoots the input signal's return to zero, these effects perfectly cancel each other, rendering the input and the summed signals identical. This characteristic is unique to a family of crossovers identified by Richard Small as "constant voltage crossovers." The first-order filter set is the only symmetric low-pass/high-pass filter pair that falls into this class.

By contrast, the summed second-order impulse response shown in Fig. 17-47 contains significant deviations from ideal.

Note that, in the summed signal (output) there is overshoot on the return to zero, followed by a delayed reaction due to the low-pass filter's delay characteristics. Viewed in the frequency domain, the second-order summed low-pass and high-pass response has a perfect null—i.e., a notch that is infinitely deep on a decibel scale—at the crossover frequency. Its phase response goes through a wrap of 360° centered at the crossover frequency.

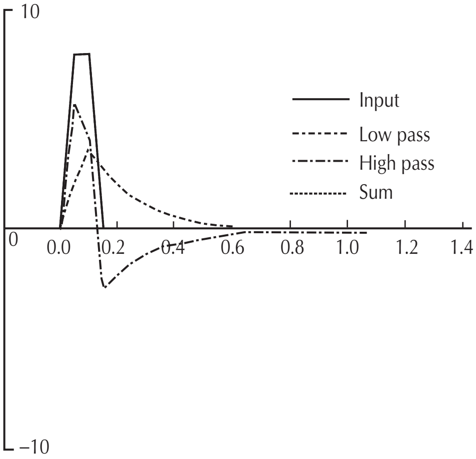


Figure 17-46. Impulse response family of first-order crossover.

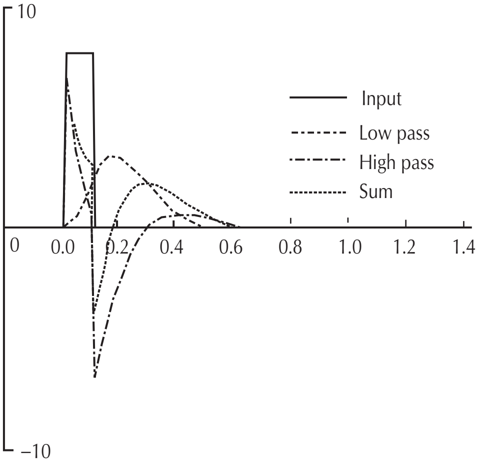


Figure 17-47. Impulse response family of second-order crossover.

Fig. 17-48 shows the impulse response family of a fourth-order Linkwitz-Reilly filter pair. It should be evident from this series of graphs that the impulse response of a loudspeaker may be compromised by the designer's choice of crossover filter topologies. Viewed in the frequency domain, the Linkwitz-Reilly filter pair exhibits ideal amplitude response (i.e., perfectly flat) through the crossover range and elsewhere, but its phase response goes through a 720° wrap through the crossover region. From what we have observed so far, it is evident that higher-order symmetric filters can introduce nonideal transient response behavior when used as crossover filters. Based on observations of the delaying effect of the low-pass filters, one might be tempted to introduce electrical delay into the high-frequency signal in an attempt to better synchronize low- and high-frequency signals. In the case of the LR filter family, such attempts will only serve to compromise the amplitude response of the loudspeaker while offering minimal improvement in the impulse response. As with crossover-frequency anomalies caused by noncoincident transducers, one way of addressing the nonideal behavior of higher-order symmetric filters is to assert that the problems are not audible. It is also possible to address transient response issues and at the same time retain a steep filter slope for one of each pair of neighboring bands in a multiway loudspeaker. Crossovers of this type are termed *constant voltage* crossovers and are discussed in Section 17.8.5.3. The simulations above are based on ideal filter behavior and ideal transducers. As one makes the simulation more realistic, accounting for the bandpass behavior of real-world transducers, the performance of all of the modeled crossovers will deteriorate, but the relative attributes of constant voltage filters remain.

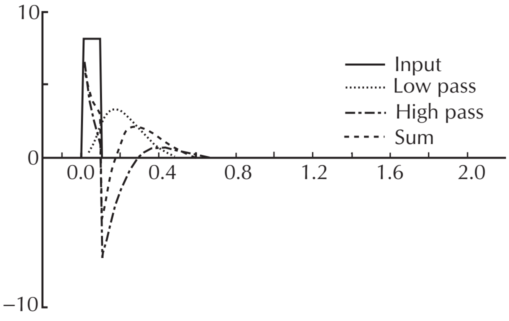


Figure 17-48. Impulse response family of fourth Linkwitz-Reilly filters.

17.8.5.4 Beyond Two-Way Systems

As the number of spectral bands in a loudspeaker increases, the issues that must be dealt with in crossover design multiply. In a system with three or more bands, at least one of the crossover filters is a bandpass, usually formed by cascading low-pass and high-pass filters of the desired characteristics. The low-pass portion of the arrangement will introduce delay in its passband, which can create misalignment between the band in question and its lower neighbor. In addition to this issue, there is also the possibility of interactions between transducers that are not neighbors in the audio spectrum (e.g., the woofer in a three-way system can contribute enough energy in the high-frequency horn's passband to make its presence known). This type of interaction is often undesirable, as it has generally deleterious effects on the response and directivity of the system.

17.8.5.5 Passive versus Active Crossovers

When designing a passive crossover—one that receives the power amplifier's output and applies appropriately filtered signals to each transducer—the designer must account for the frequency dependence of the impedances of each transducer in the system. In the case of most cone transducers, the impedance curve has a peak at the resonant frequency, above which it decreases to a minimum and then rises with frequency in similar fashion to the impedance of an inductor. This variation of impedance with frequency is often minimized through the use of a parallel, or shunt, network. Once the device's impedance has been stabilized in this manner, the actual crossover filter may be designed to drive a purely resistive load with excellent results. Active crossovers —those that divide the spectrum at line level and apply the band signals to the inputs of power amplifiers— have the advantage of the buffering effect provided by the power amplifier. Impedance-related issues are far less significant in this case, and active filters—particularly DSP-based ones—offer a number of options not readily available in passive versions. These include frequency-independent delay, all-pass filters, and dynamics processing (compression/limiting). The price that is paid in an active system is in additional channels of power amplification and wiring.

17.8.6 Acoustic Boundaries

Generally, one considers that acoustic boundaries are part of the space into which a loudspeaker is radiating. The field of architectural acoustics is largely concerned with the acoustic behaviors such boundaries cause. However, every loudspeaker has a collection of acoustic boundaries independent of the external environment in which it is operated, and these boundaries make a surprisingly large contribution to the loudspeaker's response and directivity.

Most of the boundaries associated with loudspeakers constitute reflective surfaces: enclosure walls are designed to be rigid and generally have hard surfaces. The same is true for horn surfaces. Phenomena associated with this type of surface fall into two broad categories: reflection and diffraction.

In the simplified textbook models such as a piston radiating into a half space, the infinitely large baffle on which the loudspeaker is mounted is assumed to be perfectly reflective. All of the reflections that occur at this surface will add coherently to the outgoing wave, since the source is in the same plane as the baffle. The only interfering radiation present in this model is that which is caused by the source itself, and it is this simplicity that allows a closed form solution—the piston directivity function—to yield an accurate prediction of the device's behavior.

If a hard surface is present on the front of the baffle and at right angles to it—as would be the case with room walls, for example—the wave's outgoing motion can continue no farther past this surface. Its direction is reversed due to reflection.

In a typical direct radiator loudspeaker, the wave created by a transducer expands along the front surface of the cabinet until it reaches the edges. At these edges, the support provided by the enclosure's front surface for forward motion of the wave abruptly collapses as the wave is allowed to expand rearward as well as forward. The propagation of the sound wave past this point is altered by diffraction.

Loudspeaker cabinet diffraction has not been a well-understood phenomenon until relatively recent work. The model developed by Vanderkooy shows that diffraction at an edge has strong dependence on the observation angle and that forward diffraction (in the same direction as the original outgoing wave) is inverted in polarity, whereas diffraction at angles greater than 180° (to the rear of the loudspeaker) is of the same polarity. The reader is encouraged to study Vanderkooy's work, as well as the other references, for mathematical treatments of this phenomenon.

The net effect of this diffracted energy is to introduce a set of acoustic arrivals at an observation point that follow the direct arrival in time and are reversed in polarity for positions in front of the loudspeaker. These arrivals interfere with the direct signal, with the specific effect of the interference depending on frequency, baffle size, and transducer positioning on the baffle. The result is a series of peaks and dips in the loudspeaker's response due entirely to the baffle itself.

Some effects of diffraction from panel edges are illustrated in the following graphs. On-axis response measurements were performed on a 1 inch soft dome tweeter with a 3.75 inches (95 mm) square mounting panel. Fig. 17-49 is a response measurement of the tweeter alone, suspended from a microphone stand. Fig. 17-50 is the same tweeter mounted on a thin panel approximately 19 inches (483 mm) square.

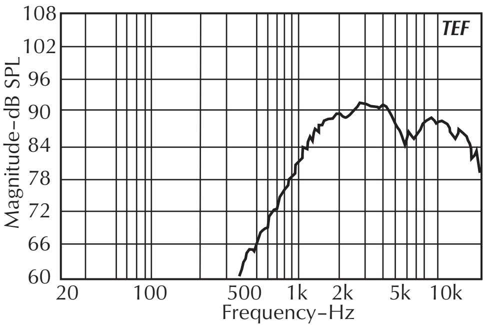


Figure 17-49. Dome tweeter on axis with no baffle.

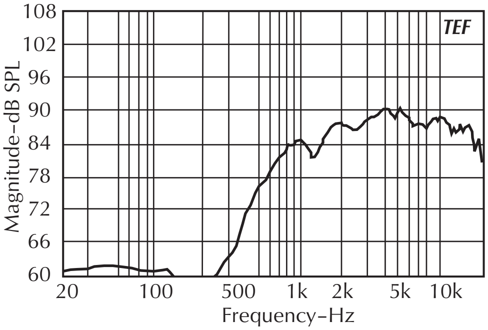


Figure 17-50. Dome tweeter on axis with 19 inch square baffle.

Note the relatively wide depression in the tweeter's response in Fig. 17-49. The center of the depression is approximately at 6.5 kHz. A diffracted arrival at a one-wavelength distance will interfere destructively with the primary wave. At 6.5 kHz, this distance is approximately 2.1 inches. This is consistent with the average distance from the center of the tweeter mounting flange to its edge. A tweeter with a round mounting flange could be expected to have a deeper, narrower notch due to reduced time smear in the diffracted arrival.

The same characteristic notch is present in Fig. 17-50, but at a much lower frequency. This is also consistent with the model of reversed-polarity forward diffraction: The notch is now centered at 1220 Hz, which has a wavelength of approximately 11 inches. The average distance from the center of the 19 inch panel to its edge matches this dimension very closely.

Fig. 17-51 is the same configuration as in Fig. 17-50, with the addition of a layer of ¾ inches (19 mm) thick foam attached at the edges of the panel. This material is relatively absorptive above 1 kHz. Its effect on the tweeter's response is most evident between 1 kHz and 3 kHz. The graphs in Figs. 17-50 and 17-51 display loudspeaker response differences that are due entirely to the boundaries formed by the speaker's baffle. The same transducer was used in each measurement.

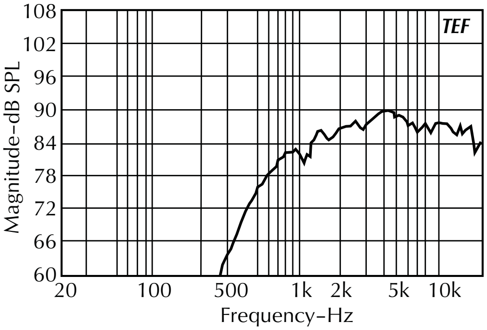


Figure 17-51. Dome tweeter on axis, mounted on 19 inch square baffle with absorption on its edge.

This brief examination of some acoustic effects due to loudspeaker boundaries is intended to provide a starting point for further study and investigation. A number of implications for loudspeaker design should be readily apparent.

Transitional points in a loudspeaker's shape (e.g., edges, slots) behave as acoustic sources. Energy arrivals from these features always follow the primary wave in time. Additionally, they can be reversed in polarity. Acoustic absorption is a useful diagnostic tool as well as a powerful design element for the loudspeaker engineer.

17.8.7 Conclusion

Loudspeaker system performance is a function of several elements, including transducer design, crossover topology, component location and orientation, and the acoustic boundaries formed by the loudspeaker's housing. Each of these elements has a major effect on the final result, and the most effective loudspeaker designs successfully address all of these areas.

17.9 Characterization of Loudspeaker Performance

17.9.1 Motivation

In considering the behavior of a loudspeaker, it stands to reason that we need performance parameters with which we can evaluate the effectiveness of a specific device for an envisioned use. There are many performance areas in which loudspeakers differ in significant ways, including on-axis response, bandwidth, directivity, distortion, and maximum acoustic output. Unfortunately, there are a number of different formats for presentation of loudspeaker performance data. Before attempting to interpret such data, it behooves us to develop some general concepts of loudspeaker performance.

The picture is made much more complicated by the fact that we hear not only the direct sound produced by a loudspeaker, but also the reflections caused by interactions between the loudspeaker and the acoustic environment in which we are listening. Different loudspeakers interact in different ways with acoustic environments, with certain types and degrees of interaction being preferable to others. For this reason, it is useful to develop a concept of loudspeaker performance that will provide a means for understanding (and, hopefully, predicting) these interactions.

17.9.2 Efficiency and Sensitivity

Since a loudspeaker converts electrical energy into acoustic energy, the concept of efficiency is relevant. As we will see, this conceptual construct, while it is a good starting point for study, has limitations in the characterization of loudspeaker performance.

Efficiency is defined as the ratio of power provided by the system output divided by the power applied to the input. As a result of conservation of energy, the efficiency of a loudspeaker (or any energy-conversion device) is always less than one. Most often, efficiency is expressed as a percentage. Typical loudspeaker efficiencies range from less than 1% in the case of some hi-fi products to approximately 25% for limited-bandwidth horn-loaded devices.

Since a loudspeaker's efficiency varies with frequency, a single number for efficiency does not generally provide adequate information for discriminating one device from another. Also, since human hearing responds to changes in acoustic pressure, the total power radiated into an acoustic space may or may not be a good indicator of what the human ear–brain perceives. Furthermore, the devices that drive loudspeakers—incorrectly called *power amplifiers*—are designed to control the voltage applied to a loudspeaker. For these reasons, the concept of a loudspeaker's functional efficiency needs to be expanded.

The parameter most often used to characterize a loudspeaker's ability to produce acoustic output is called *sensitivity*. A loudspeaker's sensitivity is the sound pressure level (SPL) produced at a reference distance with a reference electrical input signal. The most common standard is dB-SPL at 1 meter with a 1-watt input. Since a loudspeaker's impedance varies with frequency, and since a power amplifier is actually a voltage-controlled voltage source, the 1 watt figure is usually translated into an rms voltage (e.g., 2.83 Vrms into 8 Ω = 1 W). Also noteworthy is the fact that the actual measurement will generally not be accurate if it is actually carried out at a 1-meter distance, since it will not be in the loudspeaker's farfield. Instead, the testing is done at a greater distance and the results normalized to the 1-meter reference distance. The discussion that follows is based on the premise that we are interested in knowing the acoustic output characteristics of a loudspeaker with known voltages applied to its input.

17.9.3 Network Transfer Function

The performance characterization of electrical circuits is a well-developed realm and is employed as the basis for much data presented in regard to loudspeaker behavior. The impulse response, $H(t)$, and LaPlace transfer function, $L(H[t]) = F(s)$, of an electrical circuit are widely used models for the linear portion of a circuit's behavior. A large body of practical mathematics has been developed to aid in manipulating transfer functions of circuits, and the subject is covered at the undergraduate level in almost every engineering discipline.

A necessary item in applying these concepts to an electrical network is a definition of which terminals constitute the input and which will be considered the output. Using these definitions, the performance of the circuit may be modeled and/or measured and the resulting data used to evaluate the suitability of the circuit for an intended use.

17.9.4 Loudspeaker Transfer Function

Before developing these concepts further, we should recognize the importance of the correlation between the response behavior of a loudspeaker and its audible (i.e., subjectively evident) performance. The inevitable limitations in the resolution of measured loudspeaker data should ideally be determined with the capabilities and limitations of human hearing in mind. Data that is too highly resolved will reveal a number of details, or artifacts, that are not likely to be audibly significant, while insufficiently resolved data will tend to smooth over relatively serious imperfections that may easily be heard. With respect to frequency resolution, constant percentage octave (log frequency) resolution would appear to correlate best with the capabilities of human hearing. If measurements are taken so as to yield 1/6-octave resolution above 100 Hz, it is unlikely that greater resolution would reveal additional features that can be distinguished by human hearing.

The concept of frequency response (more appropriately, amplitude response) is a direct consequence of the transfer function model. This is the most familiar of the many possible ways of graphically representing portions of a transfer function. It is widely assumed that a loudspeaker may be characterized by one frequency response, usually measured at a point defined to be on axis of the loudspeaker. This assumption is incomplete: a loudspeaker has infinity of transfer functions (or, interchangeably, impulse responses), one for each point in 3D space. In the interest of compactness, we could say equivalently that a loudspeaker has a transfer function with four independent variables instead of one: where $F(s)$ is sufficient for electrical circuits, for loudspeakers the equivalent expression (in Cartesian coordinates) will be $F(s,x,y,z)$.

When we consider loudspeakers, the analogy with electrical networks is incomplete due to the nature of the device's output. Whereas a two-port electrical network has a single output, a loudspeaker radiates energy into free space in all directions. If only one listener were present, and if there were also no reflections in the acoustic environment, then the loudspeaker's response at a single point—the listening location—would be sufficient to characterize what that listener would hear. If multiple listeners and reflections are present, it is no longer sufficient to consider only the single transfer function: we need much more information.

If we limit our consideration to the far field (i.e., distances many times greater than the largest dimension of the loudspeaker), then the dependence of the transfer function on distance will be reduced, in most practical cases, to a characteristic delay of

$$\tau = \frac{r}{c} \tag{17-10}$$

where,

r is the distance from the source to the observation point,

c the phase velocity of sound in air, plus a change in acoustic pressure that is inversely proportional to distance from the source due to the inverse square law.

Both of these quantities may be assumed to be frequency-independent, although there are some exceptions.

Given this simplification, the extended transfer function of a loudspeaker may be characterized on the surface of a sphere of some arbitrarily chosen radius with the source at its center. The number of independent variables is then reduced to two, yielding a transfer function that may be represented as a function of S and two angles —i.e., $F(s,\theta,\psi)$. Even with this simplification, the number of single-point transfer functions remains uncountably infinite. Clearly, further simplification will be required if the task of measuring and describing a loudspeaker's performance is to be made practically realizable.

Currently available software for the simulation of sound system performance requires data to be presented with a fixed angular resolution. The polar coordinate system that has been adopted for this purpose is most easily described as that of a globe with the loudspeaker's axis aimed at the North pole. Typically, the plane of horizontal coverage is defined as 0° rotation angle, with the lines of constant rotation equivalent to longitude and radius angles analogous to latitude. One advantage of this approach is that data points are at maximum density near the on-axis position. There is still debate regarding the angular resolution required to show relevant details of device performance. Increments as fine as 1° have been suggested. Practically speaking, even with 10° increments, a complete set of measurements on a device with mirror-image symmetry (i.e., requiring measurements in only one quadrant) requires 172 response measurements of the device. An asymmetric device (e.g., Altec VIR, IMAX PPS, requiring two quadrants of measurement) requires 325 measurements to characterize with 10° resolution.

One possible compromise is to use one angular increment for measurements taken within the intended coverage pattern of the loudspeaker and another, broader, one for other measurements. This has the advantage of providing greater detail in the angular area in which the loudspeaker's response has the greatest audible effect. It would, however, complicate the process of interpolation that is required to approximate the response of a speaker at angles that fall between the angles at which measurements were taken. This variable resolution is unavailable in currently array prediction software.

Due to the large amount of measured data that is required to meaningfully characterize the performance of a loudspeaker, it is highly impractical, if not altogether impossible, to provide the data in a hardcopy format. In order to conveniently view loudspeaker data of this complexity, one must employ a computer program. For many years, the only available programs for this purpose were those that were primarily designed for sound system modeling and prediction. These programs have capabilities that go far beyond the display of loudspeaker data, are not optimized for that use, and are typically quite costly.

Recently, a format specifically for presentation of loudspeaker performance data, called the *common loudspeaker format*, or CLF, has been developed. This format is supported by a consortium of loudspeaker manufacturers. Due to the amount of data accommodated by the format, it is optimized for electronic, rather than hardcopy, presentation of data. It requires a data viewer program, which is available for download free at http://clfgroup.org. The displays in the CLF viewer include 3D amplitude balloons, traditional polar plots, normalized off-axis response plots, impedance versus frequency, as well as other data. Figs. 17-52 and 17-53 are screen captures of a CLF display.

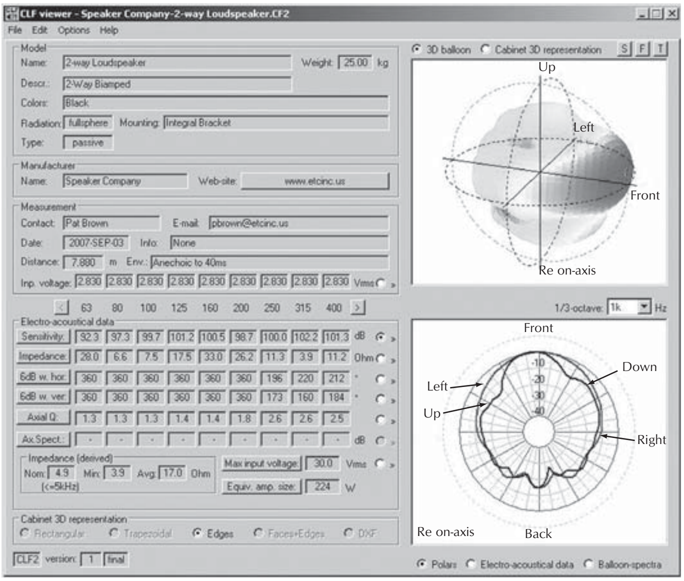
17.9.5 Impedance

The impedance of a loudspeaker is very seldom constant with respect to frequency. For this reason, the nominal impedance provided in the specification sheet—typically 4 Ω, 8 Ω, or 16 Ω—is often useless as a figure of merit. Because power amplifiers have limited ability to drive excessively low impedances and because loudspeaker cabling may have nonnegligible series resistance, it behooves the prospective purchaser of a loudspeaker to examine its impedance versus frequency curve. Of greatest interest is the minimum impedance seen in the device's bandwidth and, to a lesser extent, the frequencies and magnitudes of any peaks in the curve.

17.9.6 Distortion

The concept of a transfer function assumes a linear system. In a linear system excited by a single frequency, the output will contain only that frequency, possibly changed in amplitude and/or phase. By extension, the output of a linear system excited by a signal containing multiple frequencies will contain only those frequencies present in the original signal.

There are a number of nonlinear mechanisms in any loudspeaker. These include nonlinearities in the motor, suspension, and air (e.g., in a phasing plug or the throat of a horn). For this reason, all practical loudspeakers have nonnegligible levels of harmonic and intermodulation distortion. By comparison with modern electronic signal processing devices and amplification, loud-



CLF viewer - Speaker Company-2-way Loudspeaker.CF2

Model
Name: 2-way Loudspeaker Weight: 25.00 kg
Descr.: 2-Way Biamped
Colors: Black
Radiation: fullsphere Mounting: Integral Bracket
Type: passive

Manufacturer
Name: Speaker Company Web-site: www.etcinc.us

Measurement
Contact: Pat Brown E-mail: pbrown@etcinc.us
Date: 2007-SEP-03 Info: None
Distance: 7.880 m Env.: Anechoic to 40ms

	63	80	100	125	160	200	250	315	400	
Inp. voltage:	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	Vrms
Sensitivity:	92.3	97.3	99.7	101.2	100.5	98.7	100.0	102.2	101.3	dB ●
Impedance:	28.0	6.6	7.5	17.5	33.0	26.2	11.3	3.9	11.2	Ohm ○
6dB w. hor:	360	360	360	360	360	360	196	220	212	° ○
6dB w. ver:	360	360	360	360	360	360	173	160	184	° ○
Axial Q:	1.3	1.3	1.3	1.4	1.4	1.8	2.6	2.6	2.5	○
Ax.Spect.:	-	-	-	-	-	-	-	-	-	dB ○

Impedance (derived) Nom: 4.9 (<=5kHz) Min: 3.9 Avg: 17.0 Ohm
Max input voltage: 30.0 Vrms
Equiv. amp. size: 224 W

Cabinet 3D representation: ○ Rectangular ○ Trapezoidal ● Edges ○ Faces+Edges ○ DXF

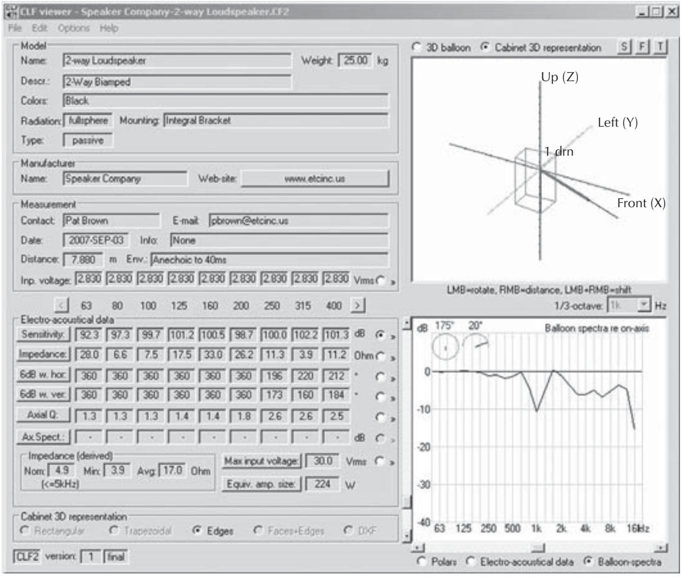
CLF2 version: 1 final

● 3D balloon ○ Cabinet 3D representation

1/3-octave: 1k Hz

● Polars ○ Electro-acoustical data ○ Balloon-spectra

Figure 17-52. CLF main window, showing balloon and polar graphics.



CLF viewer - Speaker Company-2-way Loudspeaker.CF2

Model
Name: 2-way Loudspeaker Weight: 25.00 kg
Descr.: 2-Way Biamped
Colors: Black
Radiation: fullsphere Mounting: Integral Bracket
Type: passive

Manufacturer
Name: Speaker Company Web-site: www.etcinc.us

Measurement
Contact: Pat Brown E-mail: pbrown@etcinc.us
Date: 2007-SEP-03 Info: None
Distance: 7.880 m Env.: Anechoic to 40ms

	63	80	100	125	160	200	250	315	400	
Inp. voltage:	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	2.830	Vrms
Sensitivity:	92.3	97.3	99.7	101.2	100.5	98.7	100.0	102.2	101.3	dB ●
Impedance:	28.0	6.6	7.5	17.5	33.0	26.2	11.3	3.9	11.2	Ohm ○
6dB w. hor:	360	360	360	360	360	360	196	220	212	° ○
6dB w. ver:	360	360	360	360	360	360	173	160	184	° ○
Axial Q:	1.3	1.3	1.3	1.4	1.4	1.8	2.6	2.6	2.5	○
Ax.Spect.:	-	-	-	-	-	-	-	-	-	dB ○

Impedance (derived) Nom: 4.9 (<=5kHz) Min: 3.9 Avg: 17.0 Ohm
Max input voltage: 30.0 Vrms
Equiv. amp. size: 224 W

Cabinet 3D representation: ○ Rectangular ○ Trapezoidal ● Edges ○ Faces+Edges ○ DXF

CLF2 version: 1 final

○ 3D balloon ● Cabinet 3D representation

1/3-octave: 1k Hz

○ Polars ○ Electro-acoustical data ● Balloon-spectra

Figure 17-53. CLF main window, showing cabinet drawing and normalized off-axis response.

speakers have orders of magnitude higher levels of distortion.

The degree to which loudspeaker distortion constitutes an audible problem is a matter of some controversy. Indeed, some of the more popular devices have distortion levels that are quite high, even by loudspeaker standards. Various studies over the years have established the audibility of simple harmonic distortion at levels above approximately 2%, but it is not clear that all of these studies fully accounted for the distortion present in the loudspeakers required to perform the testing. The continuing popularity of tube amplifiers among the audiophile community would tend to indicate that at least some forms of distortion are perceived as pleasing. In terms of simple harmonic distortion, it is often asserted that even-order distortion products have a more musical relationship to the fundamental, going upward in frequency in successive octave steps. Because of this, it may be true that even-order distortion products are more readily tolerated (or even preferred) by many listeners.

Because of the relatively high levels of distortion in all loudspeakers and the wide variety of ways in which a signal can be distorted, there is no consensus in the industry as to the best means for characterizing the distortion performance of a loudspeaker. The audible significance of the nonlinear distortion caused by a loudspeaker is best judged in person, and the results may or may not correlate well with commonly measured data.

17.9.7 Characterization for Design Purposes

Prior to beginning work on a new design, the designer will (or should) develop a set of performance criteria for a loudspeaker. Parameters typically considered are bandwidth, available acoustic output, directivity, and efficiency. Often, the designer must interpret subjective information provided by others and translate it into performance specifications.

The more complete the data regarding target performance, the more satisfactory the finished loudspeaker is likely to be. For this reason, targeted applications for a new design should be well understood and well defined. Possible performance definitions include characteristic isobars, lower and upper cutoff frequencies, tolerance for nonideal amplitude response (off-axis as well as on-axis), maximum acoustic output, maximum distortion level relative to fundamental, and phase versus frequency criteria. A reasonably good idea of minimum acceptable performance is also useful for purposes of cost engineering. Once the overall loudspeaker performance envelope is finalized, requirements for individual component performance can be established.

The specific measurements that should be performed on an individual component will depend on the transducer or radiator in question and on the nature of the loudspeaker of which it will become a part. Impedance versus frequency measurements are always essential. For woofers, this will allow determination of the parameters necessary for the design of an appropriate enclosure. For horn/driver combinations, the designer needs to know the frequency of mechanical resonance. It is also possible to identify internal horn reflections and diaphragm breakup problems in an impedance curve. Finally, the behavior of the component as an electrical load is required for the design of passive crossovers.

Measuring a representative set of transfer functions is an essential part of the component characterization process. What constitutes a representative set will depend on the component. A woofer will need comparatively few measurements if it is well behaved and to be used only at low frequencies, a horn will need a significantly larger number of measurements, and an array of two or more components operating over the same range of frequencies will require still more measurements. Some devices cannot be suitably characterized by a reasonable number of measurements. The measurement process will determine if a component is usable in the intended application and will aid in early prediction of the ultimate performance of the completed loudspeaker.

The extent of testing needed to evaluate the most complex component is likely to be required for the complete system. The degree to which target performance objectives have been met should be established at the prototype stage. Any necessary modifications may then be made and the system retested. This process may continue through as many iterations as necessary for the loudspeaker to perform as desired.

17.9.8 Characterization for the User

Once a loudspeaker is in production, performance data will be required for (a) giving potential buyers information for comparative purposes, and (b) use in the design of sound systems. Loudspeaker performance data provided to the sound system designer should be sufficiently comprehensive for acceptably accurate prediction of the performance of a sound system. Unfortunately, the volume of data required to fully characterize a loudspeaker is, as we have discussed, quite large. If hard copy were generated with all pertinent information, most loudspeakers intended for professional use would require a small book.

There are many ways to provide transfer function information about a loudspeaker. Keeping in mind that our extended definition of *transfer function* for a loudspeaker intrinsically includes directivity information, possible formats include:

1. Amplitude response curves calibrated to a constant level reference (e.g., dB-SPL) with a specified signal input (e.g., 2.83 Vrms) at a variety of angles. This format has the advantage of explicitly showing the direct-field response that listeners in various locations relative to the loudspeaker will hear, Fig. 17-54.
2. Amplitude response measurements as above but normalized to the response at a particular angle,

usually on axis. This is the equivalent of assuming that the on-axis response will be equalized flat, which is not always a good idea. The caveat: anomalous narrowband behavior on axis (i.e., a notch) that disappears off axis will create features (peaks) in the normalized curves that are unrepresentative of what would occur in actual use.

3. On-axis amplitude response, accompanied by polar plots at various frequencies. The common usage of only vertical and horizontal polar curves is problematic. The omission of polar curves at angles between 0° and 90° rotation leaves a lot of a loudspeaker's performance to speculation, Fig. 17-55.
4. On-axis amplitude response, accompanied by isobars at various frequencies. This format is useful for showing overall coverage behavior of a loudspeaker. Lobes, where present, are not generally revealed in isobar plots.
5. Directivity data for use with sound system modeling software. The format of this data will be dictated by the requirements of the predictive program. Some standards for the presentation of this are being discussed, but there is not at yet an industry-wide consensus on a final standard, Figs. 17-56 and 17-57.

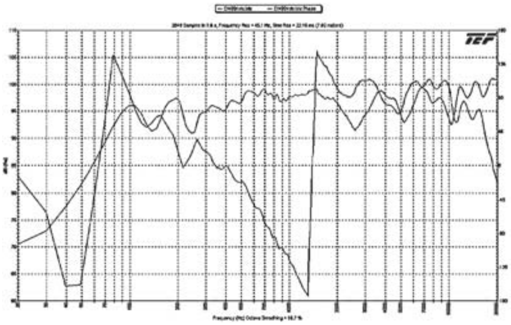


Figure 17-54. Graph of a loudspeaker's on-axis amplitude and phase response, $^1/_6$-octave smoothed. Courtesy Frazier Loudspeakers.

Of course, various combinations of the above can be provided. The formats available for the presentation of loudspeaker data continue to evolve. The availability of inexpensive mass data storage media and ever more sophisticated acoustic modeling software will continue to make the presentation of loudspeaker response and directivity information more effective and intuitive.

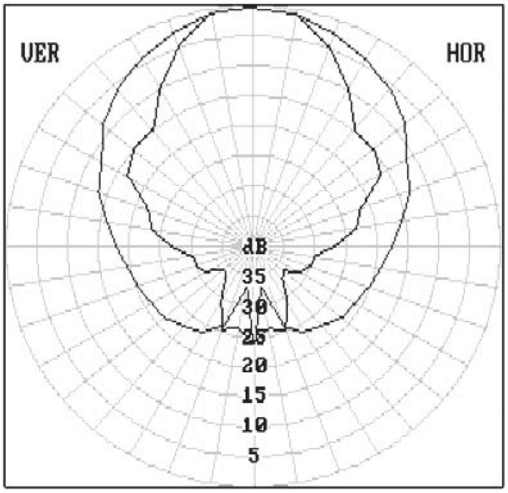


Figure 17-55. Vertical and horizontal polars, one octave averaged. Courtesy Frazier Loudspeakers.

−3 dB, −6 dB, −9 dB, coverage isobars

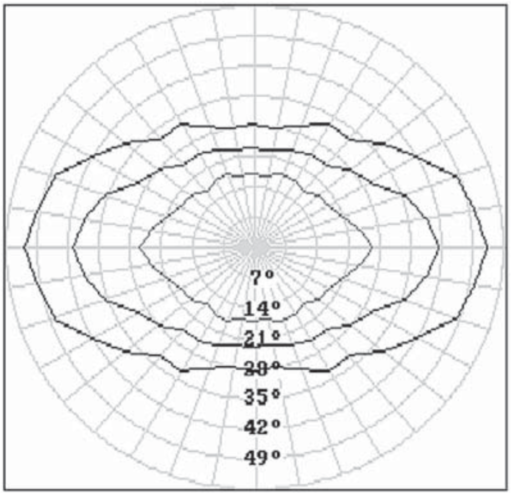


Figure 17-56. Octave band averaged isobar. Courtesy Frazier Loudspeakers.

17.10 Direct Radiation of Sound

The physics and mathematics of loudspeaker behavior are diverse and complex. In order to account for the conversion of an electrical signal to sound, one must develop both acoustic and electromechanical models. Several of these models, which are developed and presented in almost every introductory text on acoustics, are presented here without proof. The interested reader is encouraged to study the references.

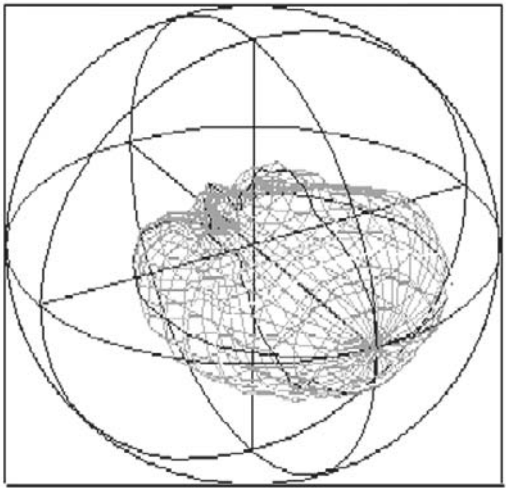


Figure 17-57. Three-dimensional representation of octave band isobar. Courtesy Frazier Loudspeakers.

17.10.1 Acoustics of Radiators

An understanding of direct sound radiation from a piston in space, a baffle, or a box can be approached by analyzing two distinct but directly related quantities, radiation resistance and directivity. Radiation resistance is the measurement of the capacity of an acoustic radiator to convert vibratory motion into sound energy. It is the ratio of pressure to the volume velocity due to the piston's motion. At high frequencies, all pistons have the same capacity per unit surface area to produce acoustic power. However, as the size of the wavelength of sound being produced approaches the size of the piston, the radiation resistance decreases as the square of frequency, i.e., at approximately 12 dB/octave.

17.10.1.1 Piston in an Infinite Baffle

A piston in a wall of infinite extent (half space) is the model most commonly employed to develop predictive equations. Even though this model is not representative of the majority of actual loudspeakers, its simplicity and mathematical manageability make it useful for instructional and comparative purposes.

A piston in an infinite baffle will see an acoustic load that depends on its size relative to a wavelength of sound at the frequency of interest. The radiation resistance, which is the part of the acoustic impedance that accounts for transmission of sound energy, is given by

$$\begin{aligned} R_f &= \rho_0 c\pi a^2[R_1(2ka)] \\ &= \rho_0 cS[R_1(2ka)] \end{aligned} \tag{17-11}$$

where,

p_0 is the equilibrium density of air,

c is the velocity of sound in air,

a is the radius of the piston,

k is the wave number, $2\pi f/c$,

$S = \pi a^2$ is the surface area of the piston,

$R_1[2ka]$ is the piston resistance function, given by

$$R_1(x) = \frac{x^2}{2 \bullet 4} - \frac{x^4}{2 \bullet 4^2 \bullet 6} + \frac{x^6}{2 \bullet 4^2 \bullet 6^2 \bullet 8} - \ldots \tag{17-12}$$

The value of the piston resistance function approaches unity for values of $2ka$ above 6. For example, in the case of a piston with an effective radius of 6 inches, the radiation resistance will be approximately constant above 1100 Hz.

The acoustic power radiated by a flat piston is given by

$$\begin{aligned} W &= \frac{R_r U_0^2}{2} \\ &= \frac{U_0^2 \rho_0 c\pi a^2 R_1[2ka]}{2} \\ &= U_0^2 \rho_0 cSR_1[2ka] \end{aligned} \tag{17-13}$$

where,

U_0 is the amplitude of the piston's velocity.

Two regimes of interest may be derived from the above equation. If we first consider $2ka < 1$—i.e., a small piston and/or low frequency—we can neglect the higher-order terms in the expression for the piston resistance function

$$R_1(x) \approx \frac{x^2}{8} \tag{17-14}$$

and the power radiated by a flat piston becomes

$$W = \frac{\rho_0 ck^2}{4\pi}(S^2 U_0^2) \tag{17-15}$$

Note that, for constant velocity amplitude, the acoustic power rises as the square of the frequency. Clearly, there must be a compensating mechanism, as a typical cone transducer has relatively flat amplitude response over this range of frequencies. This mechanism is the mechanical impedance due to the moving mass of the piston, which rises with the square of frequency. Therefore, a piston excited by a force that does not vary with frequency responds with a velocity that falls off as the square of frequency. But, in the low-frequency region, the acoustic impedance rises with the square of frequency, so the two effects effectively cancel each other over a significant range of frequencies. It is this serendipitous balance between key mechanical and acoustic parameters that makes the cone transducer an effective acoustic radiator.

The second regime is the region for which *ka* >>1 (high frequencies and/or a large piston). In this case, because the piston resistance function approaches unity, we get

$$\begin{aligned} W &\cong \frac{1}{2}\rho_0 c\pi a^2 U_0{}^2 \\ &= \frac{1}{2}\rho_0 c S U_0{}^2. \end{aligned} \qquad (17\text{-}16)$$

Note that there is no frequency dependency in the above expression: the radiation resistance of a piston approaches a constant at high frequencies. Given velocity amplitude that falls off as the square of frequency, it is clear that, in the high-frequency regime, the acoustic power radiated by a typical transducer could be expected to decrease as the fourth power of the frequency. Note also that, in the low-frequency limit, for constant velocity amplitude, radiated power goes as the square of the surface area of the piston. In the high-frequency limit, however, it goes as only the first power of the area. Therefore, all else being equal, increasing the size of the piston has a greater effect on its low-frequency output than on its capacity to radiate higher frequencies.

17.10.1.2 Piston Directivity

So far, we have examined expressions for the total power radiated by a piston. If a piston radiated identically in all directions, no further acoustic information would be needed. Since this is not the case, it is also worthwhile to consider the nature of this directivity.

The mathematical technique for deriving the piston directivity function is to consider the piston as being made up of infinitesimal differential elements, each of which contributes to the observed radiation at a point in space. These individual contributions are combined via integration to yield a value for each specific point in space.

In coming up with a manageable expression for piston directivity, one assumption must be made: the distance from the piston to the observation point is much greater than the piston's radius. The result for the pressure amplitude is

$$\rho = \frac{\rho_0 c k a U_0}{2r}\left[\frac{2J_1(\mathrm{ka}\sin\theta)}{\mathrm{ka}\sin\theta}\right] \qquad (17\text{-}17)$$

The first term in the above relationship contains the dependency of the pressure on velocity amplitude, piston size, and distance from the source. The second term, called the *piston directivity function*, is derived from a Bessel function, $J_1(x)$. The value of this function is graphed in Fig. 17-58. Note that, up to *ka* = 3.83, the value of the piston directivity function is uniformly positive. The radiation pattern of the piston will have only a single lobe under these conditions. If *ka* = 3.83, the pattern will have a null at 90° off-axis. For higher values of *ka*, this null will occur at successively smaller angles. Additionally, secondary lobes will appear outside of the main lobe, although these lobes are smaller in magnitude than the primary one. These lobes will alternate in sign: the first set will be negative, the second positive, etc.

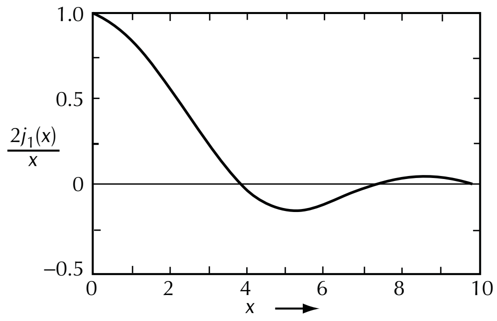


Figure 17-58. Piston directivity function.

The directivity of a real loudspeaker differs from that predicted for a rigid piston due to the fact that several of the basic assumptions in the preceding model are not fully satisfied. First, no real loudspeaker has a perfectly rigid cone or diaphragm. In the case of a cone transducer, the diaphragm is excited at its center. The excitation travels outward from the voice coil as an acoustic disturbance in the cone material. The velocity of propagation of this disturbance is always finite. At lower frequencies, this effect is negligible, but at higher frequencies not all portions of the diaphragm will vibrate in phase.

A second difference between real loudspeakers and our theoretical piston is that practical diaphragms are very seldom flat. Most often, they are in the shape of a concave cone, but convex dome shapes are also employed. In many instances, the shape of the diaphragm is chosen so as to minimize the effect of finite-velocity wave propagation in the diaphragm material on the device's on-axis response.

Generally speaking, the directivity of real-world cone or dome transducers is qualitatively similar to that of a rigid, flat piston. The nonideal behavior of real transducers can actually create beneficial effects in that the frequency at which secondary lobes appear can be higher than the theory predicts.

17.10.2 Direct Radiator Enclosure Design

A woofer is not effective as a freestanding radiator. If it were to be employed in this fashion, radiation from the rear of the diaphragm, which is out of phase with that from the front, would cause cancellation, particularly at low frequencies. Consequently, woofers are always housed. Two types of enclosures are widely used: sealed and vented.

17.10.2.1 Sealed-Box Systems

The low-frequency response of a sealed-box system may be modeled as a second-order high-pass filter. The effect of the enclosure is to add stiffness to the woofer suspension, which will modify the free-air resonant frequency of the woofer. The contribution made by the air in the enclosure to the stiffness of the diaphragm is given by

$$k_b = \frac{\rho_0 c^2 S_D{}^2}{V_B} \tag{17-18}$$

where

k_b is the box effective spring constant,
p_0 is the equilibrium density of air,
c is the speed of sound in air,
S_D is the diaphragm area,
V_B is the enclosure volume.

The spring constant of the enclosure simply adds to that of the woofer suspension, so

$$k' = k_d + k_B \tag{17-19}$$

The air mass that effectively adds to the moving mass of the diaphragm is given by

$$m_a = \frac{\rho_0 8 S a}{3\pi} \tag{17-20}$$

where,
S is the surface area of diaphragm,
a is the radius of diaphragm.

Again, this mass is additive, so

$$m' = m_d + m_a . \tag{17-21}$$

Note the dependence of the enclosure spring constant and the effective mass on two properties of air: its equilibrium density and the speed of sound. Both of these quantities are subject to significant variations with atmospheric conditions, so the degree of accuracy with which one can predict the response of an enclosure/transducer combination in actual use is intrinsically limited.

The resonant frequency of the woofer/enclosure system is given by

$$\omega_0 = \sqrt{\frac{k'}{m'}} \tag{17-22}$$

where,
$\omega_0 = 2\pi f$ is the angular resonant frequency.

The expression for the low-frequency farfield pressure response of a sealed-box woofer when driven by a constant-voltage source may be written as

$$p = \left[\frac{E_m B l \rho_0 S_d}{2\pi R_e m' r}\right]\left[\frac{\dfrac{-\omega}{\omega_0{}^2}}{1 + \dfrac{j\omega}{Q_t \omega_0} - \dfrac{\omega^2}{\omega_0{}^2}}\right] \tag{17-23}$$

where:
E_m *is* the amplitude of the applied voltage,
R_e is the dc resistance of the voice coil,
B is the flux density in the magnet gap,
l is the length of the voice coil conductor in the gap, and
r is the distance from the source to the observation point.

Q_t is given by

$$Q_t = \frac{\sqrt{k'm'}}{R_m + \dfrac{\beta^2 l^2}{R_e}} \tag{17-24}$$

where,
R_m is the woofer mechanical resistance (damping).

Note the separation of right side of the equation for pressure response into two parts. The first contains amplitude information resulting from the driving voltage, woofer parameters, and distance from the source, and the second provides frequency response information. A voltage excitation of the form of $E = E_m e^{j\omega t}$ is assumed.

From the first term in the equation, we can see several ways in which the system's output can be increased for a given distance and driving voltage:

1. Increase the flux density, B. Increasing magnet size will accomplish this up to the point at which the pole piece is saturated.
2. Increase the length l of the conductor in the gap. This will increase R_e, however, if all we do is to add turns to the voice coil.
3. Increase the diaphragm surface area, S_d. Doing so without changing the density of the material will also increase m, however.

Changing any of the above will potentially have an effect on the value of Q_t. If the total system Q has a value of 0.707, the response of the system will be maximally flat, also known as a *Butterworth* alignment. If the Q is higher than this, there will be a peak in the response just above the cutoff frequency.

If total Q is lower than 0.707, the response will fall off, or sag, in the region above cutoff, Fig. 17-59.

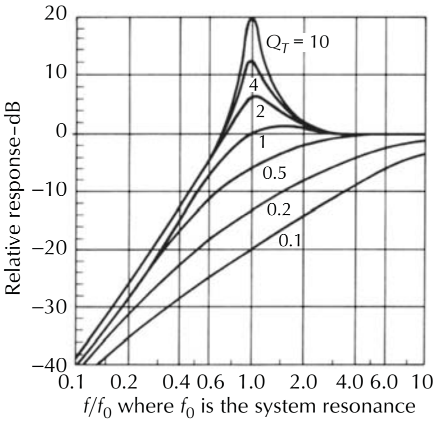


Figure 17-59. Response of closed-box system versus Q and normalized frequency relative to system resonance.

17.10.2.2 Vented Boxes

Prior to the existence of analytical models for a vented-box woofer, it was understood that an opening could be cut in a low-frequency enclosure, creating a Helmholtz resonance. The vent itself functions as an additional radiator in this case, and its radiation can add constructively to that of the woofer over a limited range of frequencies. A. N. Thiele developed the original published analytical model for vented box radiators, and his work was later supplemented by that of Richard Small.

The effect of the enclosure on the spring constant of the woofer is the same as in a sealed enclosure. The vent functions as a passive radiator coupled to the woofer cone via the air in the enclosure. In modeling the response of a vented enclosure woofer, we must account for the motion, and therefore the acoustic radiation, of the vent. The air in the vent is assumed to move as a unit to allow the mathematics to remain manageable. The following expression gives the farfield half-space acoustic pressure from a vented enclosure at low frequencies:

$$p = \left[\frac{E_m \beta l \rho_0 S_d}{2\pi r R_e m'_d r}\right] \times \tag{17-25}$$

$$\left[\frac{\left(\frac{\omega}{\omega_0}\right)^4}{\frac{\omega}{\omega_0{}^4} + \frac{1}{Q_t}\left(\frac{\omega}{\omega_0}\right)^3 - \frac{\omega_v{}^2 + \omega_d{}^2}{\omega_0{}^2}\left(\frac{\omega}{\omega_0{}^2}\right) + \frac{1}{Q_t}\left(\frac{j\omega_v{}^2}{\omega_0{}^2}\right)\left(\frac{\omega}{\omega_0}\right)}\right]$$

where,

ω is $2\pi f$,

ω_0 is $\sqrt{\omega_d \omega_v}$,

ω_v is $\sqrt{(k_v / m_v)}$.

The first portion of the right side is identical to its counterpart in the sealed-box equation. The second part describes a fourth-order high-pass filter. There are three general alignment classes for such filters: Butterworth, or maximally flat, Chebychev, or peaked; and Bessel, or maximally flat group delay.

A comparison of the attributes of sealed and vented enclosures is in order. The sealed system has the advantage of an intrinsic excursion-limiting mechanism—the addition to the woofer's spring constant due to the air in the chamber—for frequencies below the system cutoff. The vented system, on the other hand, can allow excessive woofer excursion if excited with out-of-band signals, so an electrical high-pass filter is a desirable protective element. The higher-order nature of the vented system renders it more susceptible to misalignments caused by production variations in woofer parameters and changes in atmospheric conditions, but it has the advantage of requiring less woofer excursion for a given acoustic output in the lowest portion of its usable bandwidth. In general, a sealed system will have more highly damped low-frequency transient response when compared to a vented system with the same cutoff frequency. This effect is most noticeable for frequencies in the neighborhood of the system lower cutoff frequency.

The design and modeling of vented and sealed woofer enclosures has been greatly simplified in recent years due to the ready availability of computer software developed specifically for the purpose. The following Thiele-Small (in honor of A. N. Thiele and Richard Small) loudspeaker parameters are required as minimum input to an enclosure design program:

1. Q_T is the total loudspeaker Q.
2. F_S is the free-air cone resonance of the loudspeaker.
3. V_{AS} is the equivalent volume compliance of the suspension.
4. X_{max} is the maximum linear excursion.
5. P_{max} is the maximum thermal power handling.

Most manufacturers provide the above parameters as per Audio Engineering Society (AES) recommended practice on loudspeaker specifications.

Loudspeaker enclosures can have resonances, or standing waves, caused by internal reflections. The characteristic frequencies (f_b) of these standing waves are given by:

$$f_n = \frac{c}{2}\sqrt{\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2} \tag{17-26}$$

where,

x, y, and z are the three box dimensions,

l is the designated dimension of the box,

n takes on all possible integer values ($n = 0, 1, 2, 3,...$).

If the lowest modal frequency found by setting $n = 0$ for all but the longest box dimension is above the band of use, there will be no standing waves inside the enclosure. In the more common case of a woofer being used above the first mode frequency, acoustic absorption can be added to damp these unwanted resonances. It should be understood that the addition of large amounts of damping material can have an adverse effect on the box/port tuning, so a balance must often be struck between control of standing waves and optimal low frequency response alignment.

17.10.2.3 Measurement of Thiele-Small Parameters

The most accurate way of determining f_S is by observation of a Lissajous pattern on an oscilloscope of voltage versus current with the speaker in free air. When the pattern collapses to a straight diagonal line, the phase of the impedance is zero, which indicates resonance. Once this condition has been achieved, the frequency should be measured with a frequency counter that is accurate to 0.1 Hz or better.

V_{AS} may be determined with the loudspeaker suspended in free air as follows:

1. Find total moving mass (m) by attaching an extra mass (M_X) to the cone (as close to the voice coil as possible) and observing the new resonant frequency, f_{SX}. M_X can be putty or clay; measure m_X accurately. The total mass can then be found by the equation

$$m' = \frac{M_x}{\left(\frac{f_S}{f_{SX}}\right)^2 - 1} \tag{17-27}$$

2. The suspension spring constant is then

$$k = (2\pi f_s)^2 m' \tag{17-28}$$

3. From the effective diaphragm area of the loudspeaker, S_D, V_{AS} is given by

$$V_{AS} = \frac{\rho_0 c^2 S_D{}^2}{k} \tag{17-29}$$

One means for determining the effective diaphragm area is to excite the woofer near its resonant frequency and observe its motion using a strobe light tuned almost, but not exactly, to the frequency of excitation. The resulting apparent slow motion of the woofer will allow an accurate determination of the portion of the cone that is moving.

Q_T may be found using the following procedure:

1. Determine the impedance of the woofer, Z_{max}, at its resonant frequency. This is simply the applied voltage divided by the current in the coil.
2. Identify the frequencies, f_+ and f_-, above and below f_s, respectively, at which the magnitude of the applied voltage divided by the magnitude of the current in the coil is equal to.
3. Mechanical Q is given by

$$Q_m = \left(\frac{f_0}{f_+ - f_-}\right)\sqrt{\frac{Z_{max}}{R_e}} \tag{17-30}$$

4. The total Q is then

$$Q_1 = Q_m\left(\frac{R_e}{Z_{max}}\right) \tag{17-31}$$

17.10.3 Horns

Although there are numerous mathematical treatments of horns in the texts on acoustics, they all suffer from a common set of inadequacies: the models developed in the literature account for energy transmission inside the horn, but there are no closed-form solutions to the problem of horn directivity—i.e., the behavior of a horn's radiation outside the boundaries of the horn walls, where listeners are located. Modern horn designers have been far less concerned with optimizing acoustic loading than with creating desirable directivity characteristics, and the designs have without exception been derived empirically rather than analytically.

In an exponential horn, the cross-sectional area is given by

$$S = S_0 e^{m_x} \tag{17-32}$$

where,

S_0 is the cross-sectional area at the horn's throat, or entry,

m_x is called the flare constant.

The radiation impedance of an exponential horn, assumed to be infinitely long for our purposes, is

$$Z_r = \frac{\rho_0 c}{S_0}\left[\sqrt{1-\frac{m^2c^2}{4\omega^2}}+j\frac{mc}{2\omega}\right]. \tag{17-33}$$

The first term in the brackets is the radiation resistance, and the second is the radiation reactance. Of interest is the frequency at which the value of expression inside the square root becomes zero

$$\omega_c = \frac{mc}{2} \tag{17-34}$$

or

$$f_c = \frac{mc}{4\pi} \tag{17-35}$$

This is known as the *horn cutoff frequency.* The above theory predicts that no sound will be transmitted through the horn below this frequency. Clearly, this is not the case with real horns, so the theory contains one or more assumptions that are not met in practice. Note also that the second term in the brackets, the radiation reactance, goes to zero in the high-frequency limit.

17.11 Loudspeaker Testing and Measurement

As with most other devices that transmit or process a signal containing information, measurement techniques have been developed specifically for the testing and evaluation of loudspeakers. Before the early 1980s, accurate, comprehensive testing of loudspeakers generally required expensive anechoic chambers or large outdoor spaces. Since that time, the advent of computer-based time-windowed measurements has revolutionized the field of acoustic instrumentation, particularly as regards the testing of loudspeakers.

17.11.1 Linear Transfer Function

One objective in testing a loudspeaker is to determine the linear portion of its characteristic transfer function (or, equivalently, impulse response). The most common means for acquiring this data is a spectrum analyzer. A spectrum analyzer applies a signal with known spectral content to the input of a system and processes the signal that appears at the output of the device to acquire the system's transfer function.

17.11.1.1 Spectrum Analysis Concepts

All spectrum analysis techniques are subject to a set of general constraints imposed by the mathematical relationship between time and frequency. It is useful to have a feeling for these constraints when gathering or evaluating loudspeaker data. Time and frequency are the mathematical inverses of each other. A signal that has only one frequency must exist for all time and, conversely, a signal that exists for a finite amount of time must contain multiple frequencies. A signal that exists only within a known time interval—i.e., at all times before time t_0 the value of the signal is zero and at all times after time t_1 the value is zero—can only contain frequencies given by the expression:

$$f = N\frac{1}{(t_1 - t_0)} \tag{17-36}$$

where,

N is an integer.

The frequency corresponding to $N = 1$ gives the best (lowest) frequency resolution that is possible in a test conducted for that precise time interval. All other frequencies will be integer multiples of this frequency. In order to have infinitesimally small frequency resolution (i.e., perfectly resolved frequency data), a test would have to be conducted for an infinite amount of time. It follows that all realizable response tests have a limit on their frequency resolution.

The effect of frequency resolution on a transfer function measurement is to smooth the appearance of a plot of the results, thereby possibly obscuring some of the details of the transfer function. This smoothing is present to some degree in all transfer function measurements. In the case of electronic devices, transfer functions are typically well behaved enough that the frequency resolution of a response test does not cause meaningful loss of detail. With loudspeakers, the opposite is often true: a loudspeaker's transfer function often has so much fine structure that a practical test will noticeably smooth out the peaks and dips in the speaker's response. The degree to which this fine structure is audibly significant is a matter of some controversy. As a result, there is no widespread agreement in the industry on the minimum desirable frequency resolution in testing loudspeakers.

17.11.2 Chart Recorders

Prior to the advent of computer-based measurement systems, the most commonly employed loudspeaker measurement instrumentation comprised a strip chart recorder and a signal sweep generator. The two devices are synchronized such that, for a given frequency in the sweep, the pen on the recorder is in the appropriate x (frequency) position on preprinted graph paper. The pen's y position would correspond to the amplitude of the signal received from the test microphone, and therefore, hopefully, to the amplitude response of the speaker at that frequency.

If the y amplifier is logarithmic, then the amplitude will be expressed in decibels. As common as the strip-chart measurement technique was prior to the 1980s, it had several prominent disadvantages:

1. There is no means of measuring a loudspeaker's phase response with this technique.
2. The measurement is incapable of discriminating between direct sound from the device under test and sound that is reflected from surfaces in the test environment. This necessitated the construction of very costly anechoic chambers. Even in such a chamber, the inclusion of some reflected sound in a strip-chart type measurement is unavoidable.
3. The measurement technique does not isolate the linear portion of a loudspeaker's transfer function. Distortion products are simply added to the amplitude of the loudspeaker's transfer function at the fundamental frequency that excites them.
4. There is no direct, accurate way to determine or control the frequency resolution of a strip chart measurement. Reducing pen speed and/or increasing chart (and sweep) speed have the effect of reducing frequency resolution, or smoothing, the data, but the degree to which this has taken place is not always apparent.
5. Data from this form of measurement is only generated in hard copy format.
6. This measurement technique provides no ready means to compensate for propagation delay: the time required for sound to travel from the loudspeaker to the test microphone.
7. Since there is no means for distinguishing between the output signal from the loudspeaker and ambient noise, the test environment must be quiet.

17.11.3 Real Time Analyzers

Although initially developed to measure the response of sound systems in their operating environments, real-time analysis has also been used to measure loudspeaker response in controlled environments. With this testing technique, a pink noise signal is applied to the loudspeaker. Pink noise is a random signal that contains equal energy for each unit of logarithmic frequency—e.g., for each octave or fraction thereof. The signal from the test microphone is applied to a series of bandpass filters of constant percent-octave bandwidth, each of which is tuned to a different band center, and the averaged output level of each filter is displayed, either on a CRT, LCD, or LED display. The display represents, within the limitations due to the measurement technique and the test environment, the amplitude response of the loudspeaker. Because the frequency content of a random signal has small fluctuations over time, the display may be averaged, or integrated, to produce a stable graph. Real time analysis suffers from the same general disadvantages as the chart recorder method of measurement.

17.11.4 Time-Windowed Measurements

The development of inexpensive computers has literally revolutionized the field of acoustic instrumentation. This is largely the result of the computer's ability to process and store large amounts of signal data. With a computer-based measurement system, processing and display of the data can be accomplished at any time after the raw data has been taken.

The effects of time windowing are present in any signal measurement, since the measurement must be initiated and completed in a finite amount of time (window). In digital measurement systems, however, the exact size of the time window, and therefore the resultant tradeoffs in resolution, are more directly controllable. There are two general approaches to acquiring a loudspeaker's transfer function via time-windowed measurements: measurement of the device's impulse response (time-domain measurement) or acquisition of the device's transfer function in the frequency domain.

17.11.4.1 Measurement of the Impulse Response

One form of input signal that is highly useful as an excitation for test and analysis purposes is an impulse. Mathematically, the signal is described by a Dirac delta function. Descriptively, an impulse is a voltage "spike" of very short duration and relatively large amplitude. An interesting property of an impulse is that it contains all frequencies at the same level. The impulse response of a loudspeaker, via the Fourier transform (or fast Fourier transform, FFT, as implemented in computer-based instruments), gives the speaker's transfer function. It is this equivalency via transform of the impulse response and transfer function that allows us to fully characterize a two-port device through measurement taken in only one domain or the other (time or frequency).

If a loudspeaker is excited with an impulse, the signal from a suitably well-behaved test microphone placed in front of the speaker will represent the loudspeaker's impulse response at that point. This signal may be digitized and post processed to yield the frequency-domain transfer function, as well as a number of other functions. The sampling process takes place over a fixed amount of time (the time window), and its initiation may be delayed to remove the effect of time required for sound to travel from the loudspeaker to the test microphone (propagation delay). Additionally, the length of the time window may be chosen so as to reject reflections from the room in which the measurement is being made. This is termed a *quasianechoic* measurement, and its availability has made it possible to acquire accurate direct-field response data on loudspeakers without an anechoic chamber.

The mathematics of Fourier series requires that the signal value be zero at the beginning and end of the time window. Since this condition is not generally satisfied, a window function is applied to the sampled data to force the endpoints of the window to zero. The effect of the window function is to create inaccuracies in the calculated transfer function, but these are generally much less than the spectral inaccuracies that would result from unwindowed (truncated) data. Various types of window functions may be used, including square (equivalent to unwindowed data), Gaussian, Hamming, and Hanning. Each has its advantages and drawbacks.

Among the disadvantages of impulse excitation is that of SNR. Because the impulse is of short duration, quiescent noise in the test environment can easily corrupt the test data. One means of reducing the effect of background noise is to average the results of multiple tests. If the noise is random, and therefore uncorrelated with the test signal, each doubling of the number of averages has the potential of reducing the relative noise level by 3 dB. Averaging increases the amount of time required to acquire the data.

Another disadvantage of impulse excitation is that it provides no means of identifying nonlinearities (distortion) in the loudspeaker. Distortion products appear no differently to the analyzer than the linear portion of the device's response.

17.11.4.2 Maximum Length Sequence Measurements

A variation on the method of impulse excitation is called maximum length sequence, or MLS, testing. In this form of measurement, the excitation signal is a series of pulses that repeats itself. The loudspeaker's impulse response is derived by calculating the cross-correlation between the input and output signals. This excitation signal has the advantage of producing higher average signal levels than an impulse, therefore improving the signal/noise performance of the test.

Additionally, the effects of certain types of distortion can be reduced substantially by running a series of tests employing a strategically chosen set of signal sequence lengths. To be accurate, an MLS test must be configured such that the duration of the signal sequence exceeds the length of the impulse response of the device under test. In the case of loudspeaker measurements, the impulse response of the acoustic environment must be accounted for in order to satisfy the requirement.

17.11.4.3 Other FFT-Based Measurements

Yet another variation on the FFT technique is a type of measurement that can use an arbitrary signal as the excitation signal. This form of measurement is a dual-channel FFT measurement, and the variations on this technique have several common elements. The technique involves sampling the signal at a point in the chain prior to the input of the loudspeaker (input), as well as the signal from a test microphone (output). The output may be sampled at a later time in order to account for the time required for sound to propagate from the loudspeaker to the test microphone. As with MLS testing, cross-correlation between input and output signals will yield the impulse response of the loudspeaker. It is also possible to perform an FFT on both input and output signals and obtain the transfer function of the loudspeaker by complex division.

The dual-channel approach has the advantage of allowing a wide range of signals, including music, to be employed as excitation. The commercially available implementations of this technique incorporate several refinements of the basic procedure described above, and these systems offer the possibility of measuring the response of a sound system while it is in operation.

SNR is a possible issue with this form of measurement, so averaging is generally performed to improve the accuracy of the results. Additionally, the spectrum of the input signal may not contain sufficient energy at all frequencies to sufficiently excite the system under test. For this reason, a *coherence* function is used to indicate those frequency ranges where the signal energy is insufficient to yield good results.

17.11.5 Swept Sine Measurements

Although the chart recorder is a swept sine measurement, it fails to take advantage of all the possibilities offered by the use of a sweep (also known as a *chirp*) as a test signal. Dick Heyser developed and patented a technique known as time delay spectrometry, or TDS. In TDS, the analyzer's receiving circuitry employs a bandpass filter, the center frequency of which is swept in synchronicity with the frequency of the signal applied to the loudspeaker. A delay may be applied to the sweep of the bandpass filter to account for the amount of time required for sound to propagate from the device under test to the microphone, hence the name of the technique.

The bandpass filter will reject frequencies that are displaced by some amount from its center frequency. If the receive delay for the filter is chosen appropriately, the analyzer will admit the direct signal from the loudspeaker, while simultaneously rejecting signals that have been reflected from environmental surfaces, thereby traveling a longer path and arriving later than the direct signal. The effect of this ability to reject unwanted reflections is the creation of a time window, even though the data is taken in the frequency domain. Additionally, the bandpass filter attenuates broadband noise by a much greater amount than it does the direct signal from the loudspeaker.

Due to the inherently high SNR of TDS, averaging of multiple tests is usually unnecessary. The number of samples analyzed is also not a function of the time window, as it is with an FFT-based analyzer. Furthermore, the bandpass filter removes distortion products, so TDS is intrinsically more capable of separating the linear transfer function from distortion products. It is also possible to use the technique to track specific harmonics while rejecting the fundamental.

Loudspeaker test instrumentation is more powerful and less expensive now than at any other time in the history of loudspeakers. While the current situation makes it possible to gather ever more detailed information about the behavior of loudspeakers, it is important to keep in mind the basics of instrumentation and spectrum analysis. This awareness will assist in identifying loudspeaker data that is suspect or incomplete.

Reference

1. R. H. Small, "Constant-Voltage Crossover Network Design," *Journal of the Audio Engineering Society*, Vol. 19, no. 1, January 1971.

Bibliography

A. A. Janszen, "An Electrostatic Loudspeaker Development," *J. Audio Eng. Soc.*, Vol. 3, no. 2, April 1955.

A. H. Benade, *Fundamentals of Musical Acoustics* Oxford, University Press, 1976.

A. N. Theile, "Loudspeakers in Vented Boxes: Part I & II," *J. Audio Eng. Soc.*, Vol. 19, no. 5, and 6, May and June 1971.

A. Wood, *Acoustics*, New York: Dover Publications, Inc., 1966.

Allan D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*, New York: McGraw- Hill, 1981.

B. N. Locanthi, "Application of Electric Circuit Analogies to Loudspeaker Design Problems," *J. Audio Eng. Soc.*, Vol. 19, no. 9, October 1971.

C. A. Henricksen, "Heat Transfer Mechanisms in Loudspeakers: Analysis Measurement and Design," *J. Audio Eng. Soc.*, Vol. 35, no. 10, October 1987.

C. A. Henricksen, "Phase Plug Modeling and Analysis: Radial versus Circumferential Types," presented at the 59th Convention of the AES, preprint 1328 (F-5).

C. R. Hanna and J. Slepian, "The Functions and Design of Horns for Loudspeakers," *J. Audio Eng. Soc.*, Vol. 25, no. 9, pp. 573–585, September 1977.

D. B. Keele Jr., "Low-Frequency Loudspeaker Assessment by Nearfield Sound-Pressure Measurements," *J. Audio Eng. Soc.*, Vol. 22, no. 3, pp. 154–162, April 1974.

D. B. Keele Jr., "What's So Sacred about Exponential Horns?" AES preprint 1038 (F-3).

D. B. Keele Jr., "Low-Frequency Horn Design Using Thiele/Small Driver Parameters," presented at the 57th convention of the AES, preprint no. 1250 (K-7), May 1977.

D. B. Keele Jr., "Effective Performance of Bessel Arrays," AES preprint 2846 (H1) presented at the 87th Convention of the AES, October 1989.

G. L. Augsburge,r "Electrical versus Acoustical Parameters in the Design of Loudspeaker Crossover Networks, *Journal of the Audio Engineering Society*, *V*ol. 19, no. 6, June 1971.

H. F. Olson, *Acoustical Engineering* Princeton: Van Nostrand, 1976.

H. F. Olson, *Music*, Physics and Engineering ,New York: Dover Publications.

H. Suzuki and J. Tichy, "Diffraction of Sound by a Convex or a Concave Dome in an Infinite Baffle," *Journal of the Acoustical Society of America*, preprint 0 (5), November 1981.

J. L. Bernstein, *Audio Systems*, New York: John Wiley & Sons, 1966.

J. M. Kates, "Radiation from a Dome," *J. Audio Eng. Soc.*, Vol. 24, no. 9, November 1976.

J. R. Ashley and A. L. Kaminsky, "Active and Passive Filters as Loudspeaker Crossover Networks," *J. Audio Eng. Soc.*, *V*ol. 19, no. 6, June 1971.

J. R. Ashley and M. D. Swam, "Experimental Determination of Low-Frequency Loudspeaker Parameters," *J. Audio Eng. Soc.*, Vol. 17, no. 5, October 1969.

J. R.Wright, "Fundamentals of Diffraction," *J. Audio Eng. Soc.*, May 1997.

John Borwick (Editor), *Loudspeaker and Headphone Handbook* London: Butterworths, 1988.

John Vanderkooy, "A Simple Theory of Cabinet Edge Diffraction," *J. Audio Eng. Soc.*, December 1991.

K. O. Johnson, "Single-Ended Wide-Range Electrostatic Tweeters with High Efficiency and Improved Dynamic Range," *J. Audio Eng. Soc.*, July 1964.

L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics*, New York: John Wiley & Sons, 1962.

L. L. Beranek, *Acoustics*, New York: McGraw-Hill, 1954.

N. W. McLachlan, *Loudspeakers; Theory*, *Performances*, *Testing*, *and Design*, Corrected Edition, New York: Dover Publications, 1960.

P. M. Morse, *Vibration and Sound*, New York: McGraw-Hill, 1948.

P. W. Klipsch, "A Low-Frequency Horn of Small Dimensions," *Journal of the Acoustical Society of America*, October 1941.

P. W. Klipsch, "Modulation Distortion in Loudspeakers," *J. Audio Eng. Soc.*, Vol. 17, no. 2, April 1969. (See Parts II and III, *J. Audio Eng. Soc.*, Vol. 18, no. 1, February 1970 and Vol. 20, no. 10, December 1972.)

P. W. Klipsch, "Modulation Distortion in Loudspeakers Part II," *J. Audio Eng. Soc.*, Vol. 18, no. 1, February 1970.

R. F. Allison, "The Influence of Room Boundaries on Loudspeaker Power Output," *J. Audio Eng. Soc.*, Vol. 22, no. 5, June 1974.

R. G. Brown, *Lines Waves and Antennas*, New York: The Ronald Press Co., 1961.

R. H. Small, "Direct-Radiator Loudspeaker System Analysis," *J. Audio Eng. Soc.*, Vol. 20, no. 5, pp. 383–395, June 1972.

R. H. Small, "Closed Box Loudspeaker Systems; Part I: Analysis," *J. Audio Eng. Soc.*, Vol. 20, no. 10, pp. 798–808, December 1972.

R. H. Small, “Vented Box Loudspeaker Systems; Part I II III and IV,” *J. Audio Eng. Soc.*, Vol. 15, nos. 5, 6, 7, and 8; June, July, August, and September 1973.

R. H. Small, “Phase and Delay Distortion in Multiple-Driver Loudspeaker Systems,” *J. Audio Eng. Soc.*, Vol. 19, no. 1, January 1971.

R. Heyser, “Loudspeaker Phase Characteristics and Time Delay Distortion Part 1,” *J. Audio Eng. Soc.*, Vol. 17, no. 1, p. 30, January 1969.

R. Heyser, “Loudspeaker Phase Characteristics and Time Delay Distortion Part 2,” *J. Audio Eng. Soc.*, Vol. 17, no. 2, p. 130, April 1969.

S. H. Linkwitz, “Active Crossover Networks for Non-Coincident Drivers,” *J. Audio Eng. Soc.*, Vol. 24, no. 1, p. 2, January/February 1976.

S. Ishii and K. Takahashi, “Design of Linear Phase Multi-Way Loudspeaker System,” AES 52nd Convention, October 1975, preprint 1059.

This page intentionally left blank

Chapter 18

Loudspeaker Cluster Design

by Ralph Heinz

18.1 Why Array? 647
18.2 Array Problems and Partial Solutions: A Condensed History 647
18.3 Conventional Array Shortcomings 647
18.4 Conventional Array Shortcoming Analysis 649
18.5 Coincident Acoustical Centers: A Practical Approach 650
  18.5.1 TRAP Horns: A New Approach 650
  18.5.2 TRAP Performance 651
18.6 18.6 Low Frequency Arrays: Beneficial Interference 652
  18.6.1 Horizontal Woofer Arrays: Maintaining Wide Dispersion 652
  18.6.2 Vertical Woofer Arrays 653
    18.6.2.1 Directivity at Frequencies Where Size Makes Horns Impractical 653
18.7 Line Arrays and Digitally-Steerable Loudspeaker Column Arrays 653
  18.7.1 What Affects Intelligibility 654
  18.7.2 Measuring Intelligibility 654
    18.7.2.1 Subjectively 654
    18.7.2.2 Objectively 654
  18.7.3 Architecture and Room Acoustics. 655
    18.7.3.1 Reverberation 655
  18.7.4 Line Arrays 655
    18.7.4.1 Digitally Steered Column Arrays 656
    18.7.4.2 Controlling High Frequency Beaming 656
17.7.4.3 Beam Steering 656
  18.7.5 DSP-Driven Vertical Arrays 657
    18.7.5.1 Acoustical, Electronic & Mechanical Considerations 657
    18.7.5.2 Point Source Interactions 658
17.7.5.3 Doublet Source Directivity 658
    18.7.5.4 Array Height versus Wavelength (λ) 659
    18.7.5.5 Inter-Driver Spacing versus Wavelength (λ) 659
  18.7.6 Multichannel DSP Can Control Array Height 661
  18.7.7 Steerable Arrays May Look Like Columns But They are not. 661
    18.7.7.1 Beam-Steering: Further Proof that Everything Old is New Again 662
    18.7.7.2 DSP-Driven Arrays Solve Both Acoustical and Architectural Problems 662
17.7.7.3 Variable Q 662
17.7.7.4 Consistent Q with Frequency 662
17.7.7.5 Ability to Steer the Acoustic Beam Independently of the Enclosure Mounting Angle 662
17.7.7.6 Design Criteria: Meeting Application Challenges 662
17.7.7.7 Horizontal Directivity is Determined by the Array Elements 663
17.7.7.8 Steering is Simple—Just Progressively Delay Drivers 663
    18.7.7.9 BeamWare: The Software That Controls Iconyx Linear Array Systems 663

This page intentionally left blank

18.1 Why Array?

For the purposes of this discussion we can define a loudspeaker array *as a group of two or more full-range loudspeaker systems, arranged so their enclosures are in contact*. System designers use arrays of multiple enclosures when a single enclosure cannot produce adequate sound pressure levels, when a single enclosure cannot cover the entire listening area, or both. These problems can also be dealt with by distributing single loudspeaker systems around the listening area, but most designers prefer to use arrays whenever possible because it is easier to maintain intelligibility using a sound source that approximates a point source than by using many widely separated sources.

18.2 Array Problems and Partial Solutions: A Condensed History

First-generation portable sound systems designed for music used a very primitive form of array: they simply piled up lots of rectangular full range speaker systems together, with all sources aimed in the same direction, in order to produce the desired SPL. This type of array produced substantial interference, because each listener heard the output of several speakers, each at a different distance. The difference in arrival times produced peaks and nulls in the acoustic pressure wave at each location, and these reinforcements and cancellations varied in frequency depending on the distances involved. So although the system produced the desired SPL, the frequency response was very inconsistent across the coverage area. Even where adequate high frequency energy was available, intelligibility was compromised by multiple arrivals at each listening location.

Second-generation systems incorporated compression drivers and horn-loading techniques derived from cinema sound reinforcement and used for large-scale speech-only systems (the original meaning of public address). When two or three of these horns were incorporated in a single enclosure with trapezoidal sides that splayed the horns away from each other, the first arrayable systems were introduced to the marketplace. These products promised to eliminate lobing and dead spots (peaks and nulls) and to drastically reduce comb filtering (interference). They did improve performance over the stack of rectangular enclosures loaded mainly with direct radiating cones. But frequency response across the coverage area remained inconsistent. In addition to the midrange and high frequency variations across the coverage area of the array, low frequency output varied from the front to the rear and side to side. Low frequency energy was focused along the longitudinal axis of the array and close to it, producing a "power alley" that gave the seats with the best views the worst sound, Fig. 18-1.

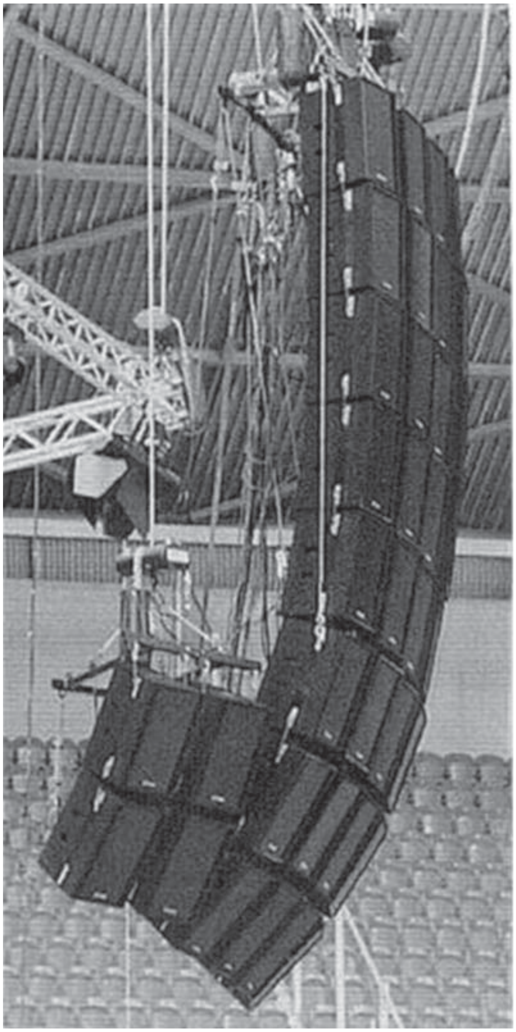


Figure 18-1. A typical second generation loudspeaker cluster. Even when a single enclosure is designed to resemble a point source, multiple enclosures will always interfere with each other when connected to a coherent audio signal.

18.3 Conventional Array Shortcomings

As we said in the first paragraph, the performance advantages of the array (whether horizontal or vertical) derive from its ability to approximate a perfect acoustical point source. But even the smallest arrays typically including three or more loudspeaker enclosures, each with two or three separate acoustic centers of its own. It's easy to appreciate that getting all those discrete sources to behave like a theoretical point source is difficult in practice. Signal processing solutions attempt to compensate for the difference between theory and reality by sacrificing the coherency of the electronic signal. They apply frequency shading and/or micro-delays to the signals sent to different enclosures, in order to ameliorate the acoustic problems. These approaches are costly, complicated and often meet with limited success.

A rigorous analysis of the acoustical physics can point the way toward a practical, physical solution. First, consider what is probably the most common arrayable system in use today: 60° × 40° horns in enclosures with 15° trapezoidal sides, Fig. 18-2.

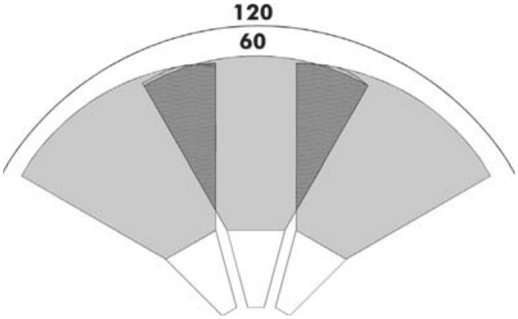


Figure 18-2. A very common array uses three 60° × 40° horns in enclosures with 15° trapezoidal sides; tightpacked, this array produces substantial overlap and interference between adjacent horns.

Tight-packing three of these systems with their 15° sides touching produces a 30° splay between the horns, for a total included angle of 120°. At first glance, this seems like an ideal alignment. But the EASE interference predictions in Fig. 18-3 show the familiar and clearly audible problems with this configuration: significant interference above 1 kHz, with variations of 8 dB – 9 dB depending on the angle. On axis, there is about 10 dB of gain at frequencies below 1 kHz. Where maximum SPL is the main consideration, this type of array will deliver acceptable performance. When the front-of-house mix position can be located on the axis of left and right arrays, they can usually be tweaked to deliver acceptable reproduction in this limited area. Other areas of the house, including the high roller seats up front, will suffer.

The interference patterns displayed in Fig. 18-3 can be reduced by widening the splay between cabinets to 30°, as illustrated in Fig. 18-4. This array will not look as pretty as the first, but it does have much more even response across the coverage area, Fig. 18-5. At 2 kHz and 4 kHz, the individual horns are clearly discernible in the ALS-1 predictions. Also note that the seams between the horns become deeper with increasing frequency.

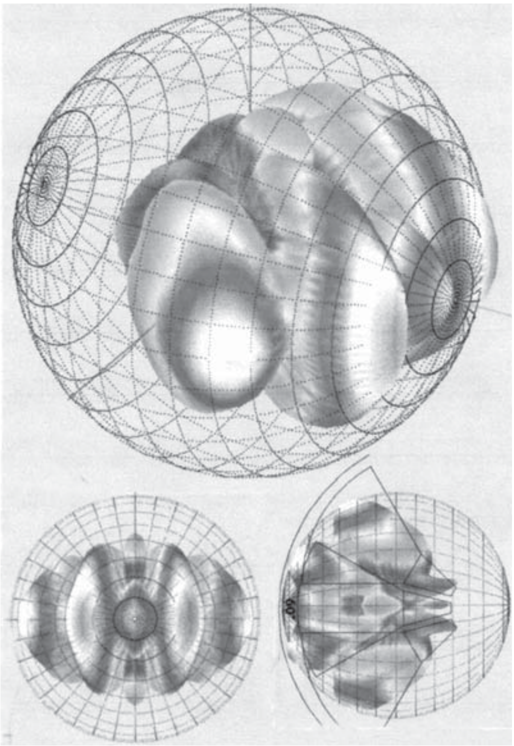


Figure 18-3. The interference patterns shown above were produced by tight-packing three arrayable loudspeakers using 60° × 40° constant directivity horns in enclosures with 15° trapezoidal sides. While this is an improvement over a pile of direct radiating transducers, it is far from the ideal point source array.

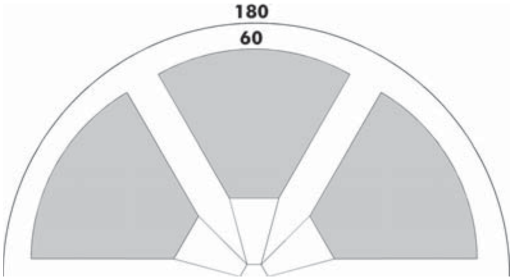


Figure 18-4. Widening the splay between horns reduces interference and widens the coverage angle to 180°, but reduces forward gain. As always, energy is conserved.

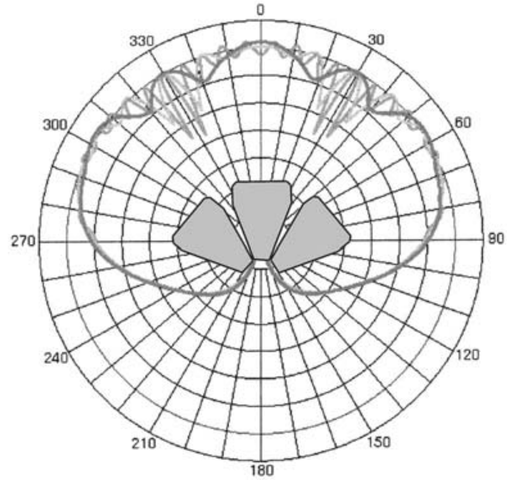


Figure 18-5. ALS-1 interference predictions for a wider splay show reduced interference, but the three horns are clearly apparent at higher frequencies.

Fig. 18-6 shows why there will always be interference with conventional horn arrays (whether they are enclosed in arrayable cabinets with trapezoidal sides or mounted in free air). As the wavefronts radiate from points of origin that are separated in space, they will always create some interference at the coverage boundaries.

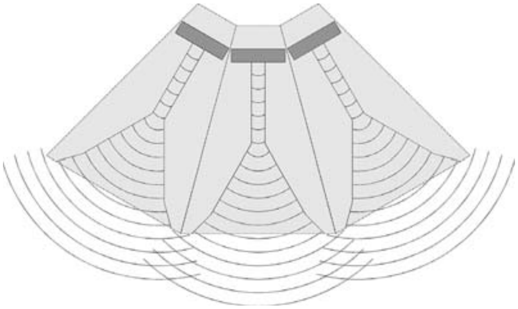


Figure 18-6. The acoustic pressure wave expands as a sphere, and multiple spherical sections will always overlap unless they originate from a common center.

18.4 Conventional Array Shortcoming Analysis

For an array in far field, dependence on angle is

$$SPL(\theta) = 10\log P_0^{\ 2}\text{ dB} \tag{18-1}$$

For a distance to the listening area very much larger than the array dimensions, let the sound pressure P be the real part of

$$P(\theta) = A_i(\theta)^{j(\omega\tau - kSi)} \tag{18-2}$$

where,

P is the sound pressure,

ω is the angular frequency,

$A_i(\theta)$ is a function of the angle between the array longitudinal axis and the direction of the distant listening point. It gives the ratio of the sound pressure due to the source as a ratio of its on-axis value at the same distance.

For the ith source shown in Fig. 18-7, assuming identical sources, the pressure contribution is given by:

$$P_i = A_i(\theta)^{j(\omega\tau - kSi)} \tag{18-3}$$

where,

k is $2\pi/\lambda = 2\,\lambda fc$,

λ is the wavelength,

f is the frequency,

c is the speed of sound,

S_i is the distance by which the path length from the i^{th} source to the distant point exceeds the distance from the origin to that point.

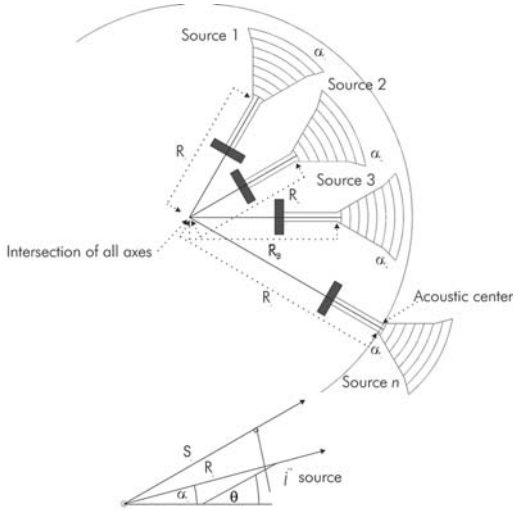


Figure 18-7. For a circular arc array, the additional path length Sj is as shown.

For an array of n sources, the total pressure P is given by:

$$P(\theta) = \sum_{i=1}^{n} A_i q e^{j\omega\tau - kSi} \qquad (18\text{-}4)$$

$$= \varepsilon^{j\omega\tau} \sum_{i=1} A(\theta) \varepsilon^{j\omega\tau kSi}$$

The square of the pressure amplitude is given by:

$$P_0^{\,2}(\theta) = \left[\sum_{i=1}^{n} A_i \theta k S_i \right]^2 + [A_i(q\theta)\sin(kS_i)]^2 \qquad (18\text{-}5)$$

where,
$A_i(\theta)$ is $A_i(\theta - \alpha_i)$.

For a circular arc array, the additional path length S_i as shown in Fig. 18-7, for the ith source at radius R and angle α is given by:

$$-S_i(\theta) = R_i \cos(\theta - \alpha_i) \qquad (18\text{-}6)$$

Therefore, the smaller R_i is, the smaller the S_i differences, and the less the interference between sources. Ideally, $R = 0$ for all sources. As R approaches 0, the interference will become less audible and frequency response across the array's intended coverage area will become more uniform.

18.5 Coincident Acoustical Centers: A Practical Approach

Clearly, the ideal solution is to collocate all the acoustic points of origin, as shown in Fig. 18-8. We could achieve this by stacking the horns vertically, but this would solve the problem in the horizontal plane by creating a worse situation in the vertical (front to back) direction. Fig. 18-9 shows a more realistic approximation that takes into account the physical constraints of loudspeaker design (the dimensions of the transducers, horns, enclosure walls, etc.). Because the acoustic sources are real physical objects, we cannot reduce R_i to 0. But we can get close enough to make measurable, audible improvements in the performance of the multi-enclosure array.

18.5.1 TRAP Horns: A New Approach

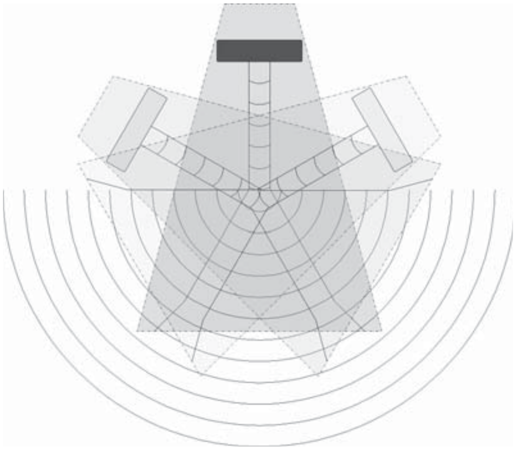


Figure 18-8. The acoustic ideal—colocating the acoustic centers of all horns is not a practical possibility.

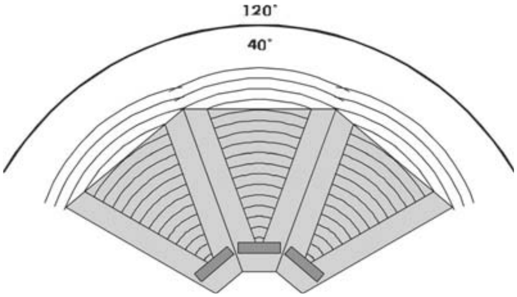


Figure 18-9. Because drivers and enclosures are physical objects, the acoustic centers of TRAP horns are not perfectly coincident□but they are close enough to achieve measurable and audible reductions in interference.

Fig. 18-9 implies that the way to minimize R_i—and the resultant interference—is to move the acoustic centers as far to the rear of the enclosure as possible. We can attempt to minimize the size of the drivers, for instance by using high-output magnetic materials such as neodymium. But the biggest obstacle to coincident acoustic centers is the horn itself. This is because typical constant directivity horns exhibit astigmatism: their apparent points of origin are different in the horizontal and vertical planes. In order to create a wider coverage pattern in the horizontal plane, the apparent apex is moved forward, while the vertical apex is farther to the rear because its coverage pattern is usually narrower. This is certainly the case with the most popular horn patterns in use today: 60° × 40° and 90° × 40°. One approach to a solution, then, is to rotate the horn and use the vertical apex of the horn in the horizontal plane. By doing so, we are effectively moving the acoustic center as far to the rear of the cabinet as possible. This technique when combined with cabinet design that minimizes the space between adjacent drivers in an array, while matching the trapezoidal sides with the opening angle of the horn, creates a system capable of minimal interference in the frequency range where the horn is effective. This forms the basis for what I call the True Array Principle by Renkus Heinz.

Subsequent refinements to the horn flare itself have been awarded U.S. Patent #5,750,943. This Arrayguide topology goes even farther in locating the apparent acoustic origin toward the rear of the enclosure. To repeat, moving the acoustic centers to the rear minimizes *R*, the distance between acoustic points of origin within the array, and the resulting interference between array elements.

Fig. 18-10 shows the ALS-1 predictions for the first generation of TRAP horns. It is clear that interference has almost disappeared.

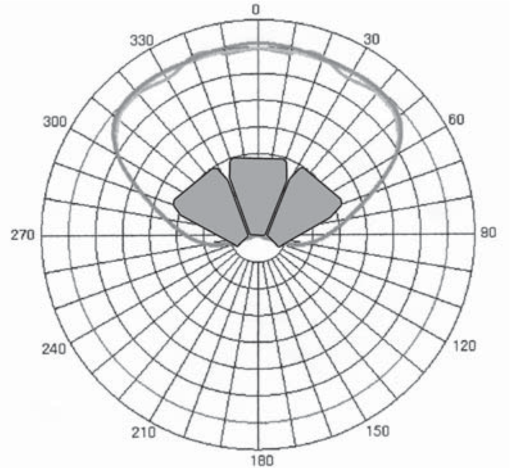


Figure 18-10. TRAP design produces truly arrayable systems with minimal destructive interference in the horns' passband.

Fig. 18-11 shows measured EASE data for a three-wide array of TRAP40 enclosures. Frequency response is consistent in both vertical and horizontal planes within ±4 dB. This is an out of the box array, using no frequency shading or micro-delay to improve performance. Measured results don't track the predictions 100% because the actual pattern of the horns varies somewhat with frequency: first generation TRAP horns maintain nominal coverage ±10° from 1 kHz to 4 kHz.

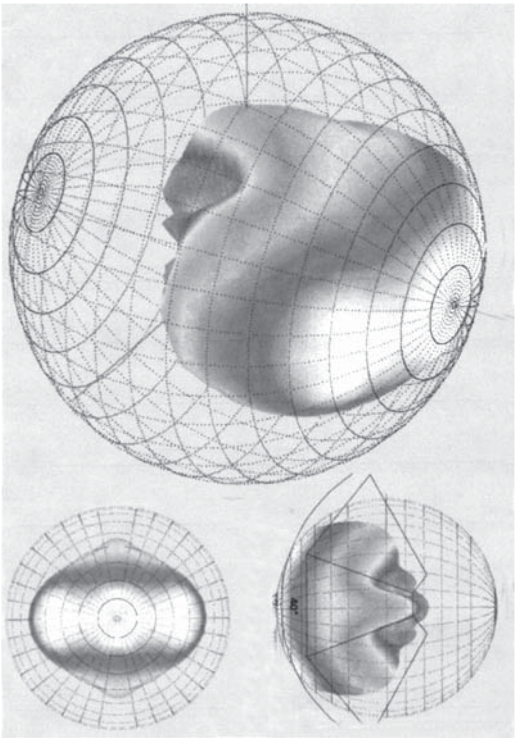


Figure 18-11. The TRAP array produces almost no measurable interference from a tight-packed three-wide cluster. This is because the three spherical wave-fronts produced by the three horns originate from a common acoustical center. Therefore they behave as a single acoustic unit, without overlap or interference.

18.5.2 TRAP Performance

Systems based on the True Array Principle can extend pattern bandwidth (the frequency range over which coverage varies less than ±5°) down to the frequency at which mutual coupling between adjacent cabinets ceases. TRAP systems are designed so that the enclosures provide optimum splay angles of 40° between the horns: the trapezoidal sides are therefore steeper than many other designs at 20° per side. The combination of symmetrical horns and steeper sidewall angles maintains coincident acoustic centers for all the elements in the array.

Note that moving the horizontal apex to the same location as the vertical results in a symmetrical 40° × 40° coverage pattern. This in turn requires the use of four enclosures to cover 160° with almost no variation in frequency response in the horizontal (side to side) plane. With 60° × 40° cabinets we could deliver sound to 180° of coverage, albeit with some quite audible variations.

There are other commercially available systems offering similar array performance to that described above. The ARC's system from French loudspeaker

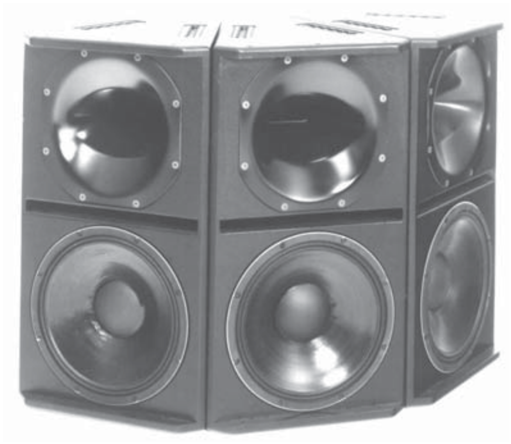


Figure 18-12. TRAP arrays can be quite small; however, the size of the horns will determine the lower frequency limit at which the True Array Principle ceases to operate.

manufacturer L-Acoustic uses a type of path length equalizer to force the emerging wavefront to conform to the opening angle of their horn and also puts the acoustic center behind the cabinet. In the case of ARC, the cabinet's trapezoidal side walls also serve as the waveguide for the high frequencies. As the waveguides opening angle matches that of the cabinet, this is certainly an elegant solution to creating minimum interference arrays at the frequency where the horn is effective.

In the KF900 series from EAW, simple phase horns for the mid and high frequencies put the acoustic center as close to the rear of the cabinet as possible, while their opening angles also match the trapezoidal sides of the enclosure. The relatively large size of the KF900 series enclosures and horns brings minimum interference performance to frequencies lower than those based on smaller waveguides. Remember, that this technique for minimum interference arrays, including the True Array Principle, only holds true for those frequencies where the horn is effective.

18.6 18.6 Low Frequency Arrays: Beneficial Interference

In the preceding paragraphs, I outlined the parameters necessary to minimize destructive acoustic interference between adjacent cabinets or horns in an array. But these techniques are only beneficial at the frequencies where the horns are effective. Yet these very systems or horns are used at frequencies well below their directivity cutoff and lower, down to frequencies where the woofers piston size offers no directional control at all.

18.6.1 Horizontal Woofer Arrays: Maintaining Wide Dispersion

For our first example, let's look at the additional problems and opportunities we create when arraying small (12 inch woofer, 1 inch compression driver) full range loudspeaker enclosures as in Fig. 18-12. For a full range array module, there are three frequency zones that exhibit different wavelength related behavior. At the lowest frequencies, or longest wavelengths, these modules exhibit only beneficial interference or mutual coupling. Each additional module creates additional on axis acoustic output. The opportunity here, is that less equalization is required to make the array's frequency response flat down to these lower frequencies as compared to a single cabinet.

A potential problem is created when the array becomes too wide however. Four or five element arrays are wide enough as to become quite directional in the forward plane at those lower frequencies (20 Hz to roughly 500 Hz or more, dependant on the module). Without a signal processing scheme, this array cannot be equalized to have the same frequency response through out it's intended coverage. It will sound boomy in the middle and thin at it's coverage extremes. A solution is to taper the length of the array in the horizontal plane in order to maximize horizontal dispersion of the lower frequencies. The entire array can be used for the lowest frequencies as the wavelengths are longest (20 Hz up to about 200 Hz), but at higher frequencies, as the wavelengths get shorter, the array length must also get shorter to maintain wide dispersion. This is achieved by low passing the outermost woofers of the array, such that only two or three max woofers are used at frequencies higher than this.

The second frequency zone that can be problematic in arrays based on full range modules, occurs at wavelengths where cabinet spacing no longer supports mutual coupling, and the horn has yet to attain it's directivity cutoff. This typically applies to a small half octave range where adjacent cabinet spacing approaches a wavelength. Here we observe combinations of destructive and constructive interference at various observation points around the array intended coverage, causing frequency response variations greater than ±6 dB. Fortunately there is a signal processing technique that can minimize this effect. By simply notching this frequency range from every other cabinet with a cut equal to the greatest amount a variance (typically 6 dB of attenuation), and width equal to the bandwidth of the aberrations (typically half an octave), the frequency response variations throughout the arrays coverage can minimized.

The third frequency zone of wavelength related behavior for arrays based on full range modules, is then at the frequencies above which the horn is effective. Let us assume that the horns depicted in Fig. 18-12 place the acoustic center towards the rear of the cabinets, and that their opening angle also matches that of the trapezoidal sides of the cabinet. Based on these assumptions, the array performance will exhibit minimum interference for frequencies above 1–2 kHz which happens to be the effective directivity cutoff of the horn. Each additional module simply adds additional coverage to the array.

18.6.2 Vertical Woofer Arrays

18.6.2.1 Directivity at Frequencies Where Size Makes Horns Impractical

Beneficial destructive interference sounds like an oxymoron, but there are several commercially available woofer arrays that take advantage of this very technique. By applying the fundamental physics described by Harry Olson, directional woofer arrays are now available that out perform large woofer horns.

When two point sources are superimposed on one another, their outputs simply add up in all directions. As the two point sources are spread apart, the output diminishes along the plane of separation due to phase cancellation. At exactly ½ wavelength, a pure null occurs, and we achieve the classic figure eight, dipole polar pattern. The current commercially available systems take advantage of this phenomena, directivity through off axis attenuation, by placing woofers in a vertical array and spacing them to create this dipolar pattern at frequencies below which horns become too large.

Fig. 18-13 is an example of one such array. Termed *Tri-Pola*r by it's designer Vance Breshears, it uses the vertical spacing between the three woofers with appropriate signal processing to maintain consistent low frequency pattern control from 400 Hz down to below 100 Hz. One of the first systems available was developed by Craig Janssen, termed *Tuned Dipolar*, it uses two separate arrays. With drivers, spacing and signal processing appropriate for their respective passbands Tuned Dipolar offer exceptional low frequency pattern control over an extended bandwidth. Even subwoofers are now benefitting from this type of technology. Meyer Sound is achieving cardioid patterns at lowest frequencies from its PSW-6, providing significant attenuation of those frequencies directly behind the enclosures.

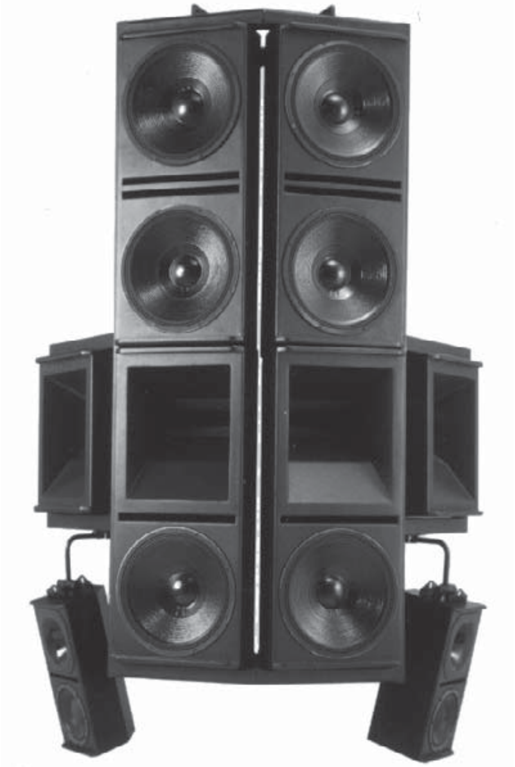


Figure 18-13. Reference Point Array using four 40° × 40° mid-high enclosures and six low frequency modules in Tri-Polar configuration for vertical pattern control, along with appropriate small full range systems for downfill.

18.7 Line Arrays and Digitally-Steerable Loudspeaker Column Arrays

For the communication between a source and a listener to be effective, it is important that the listener receive and comprehend the message. In large spaces where people gather, including auditoria, houses of worship, sports venues, transit terminals and classrooms, often the acoustic requirements that enable effective speech are in conflict with the architectural needs of the spaces. When the acoustics of a venue cannot be altered to enable effective speech communication, designing a sound reinforcement system to do so, can be a challenge. Recent advances in efficient amplification and digital signal processing have enabled a new class of loudspeaker; the *digitally steerable column* or *line array* as its often called. The acoustical and architectural benefits of these loudspeakers for sound reinforcement in highly reverberant or reflective environments will be shown.

We will discuss effective communications and define intelligibility and how to measure it both subjectively and objectively. We will look at architecture and acoustics and at reverberation and its effect on intelligibility in large public spaces. Finally we'll look at digitally steerable column arrays, their design considerations, and their performance and benefits when used in large reverberant spaces.

Some of the basic principles involved in voice communications are:

- In voice communications intelligibility is the capability of being understood.
- It assumes the existence of a communication process between a talker and a listener, or between a source and a listener.
- For the conveyance of meaning, the English language is highly dependant upon the effective receipt and comprehension of consonants. This is how we differentiate words based on similar vowels. For example, Zoo, Two, New.
- In terms of frequency response, speech ranges between 100 Hz and 8 kHz, with maximum energy around 250 Hz.
- In speech, the frequency range that conveys the most consonant information is the octave around 2 KHz.

18.7.1 What Affects Intelligibility

Major Influences that affect intelligibility are:

- Elocution and pronunciation of the talker. It's hard to understand someone who mumbles under any condition.
- Hearing acuity of listener. An often overlooked influence, those with a hearing loss have trouble understanding what's being said.
- SNR. We've all been places where it was so noisy we couldn't understand what was being said.
- Direct to reverberant ratio. The higher the reverberation level, the more difficult it is to understand what's being said.
- Directivity of the loudspeaker or loudspeakers. Highly directional loudspeakers direct more of the sound onto the audience and less onto the reflective walls and ceilings.
- The number of loudspeakers. Larger numbers of loudspeakers translate into more acoustic energy being transmitted into the room and higher reverberation levels.
- Reverberation time. The longer the reverberation time, the more likely it will interfere with intelligibility.
- Distance of source to listener, The closer the listener is to the loudspeaker, the less likely reverberation will interfere.

Secondary Influences are:

- Gender of talker.
- Microphone technique.
- Vocabulary and context of speech information.
- Direction of main sound to listener and/or direction of reflections and echoes.
- System fidelity, equalization, and distortion.
- Uniformity of coverage.

18.7.2 Measuring Intelligibility

18.7.2.1 Subjectively

Statistical tests with trained talkers and listeners can be the most reliable metric for determining the intelligibility of a system. To ensure that all speech sounds are represented in a test, Phonemically Balanced (PB) word lists are commonly used. These word list can be a long as 1000 words. Tests using nonsense syllables or logatoms, and Modified Rhyme Tests are also used. These tests are very time consuming and are difficult to set up.

18.7.2.2 Objectively

Articulation Index. Articulation Index or AI was one of the first attempts to quantify intelligibility with measurements. AI is primarily concerned with the affect of noise on speech. The index ranges from 0 to 1 with 0 representing no intelligibility.

%ALcons. %ALcons or the articulation loss of consonants was developed by Peutz in Holland during the 1970's. %ALcons takes both noise and reverberation into account and is based the importance of the octave around 2000 Hz in conveying consonant information. %ALcons uses a scale running downwards from 0 where 0 is perfect intelligibility, or 0% articulation loss.

Although Peutz used 2000 Hz as the center frequency and 2000 Hz is still the European standard, many acousticians in the USA prefer using 1000 Hz. As a general rule, %ALcons calculated at 1000 Hz show a higher articulation loss than ones calculated at 2000 Hz.

STI. STI or Speech Transmission Index considers the source/room/listener as a transmission channel and measures the reduction in modulation depth of a specialized test signal which replicates the burst nature of real speech. The STI scale ranges from 0 to 1, where 1 represents perfect intelligibility. STI is considered the most accurate of the intelligibility measures.

Evaluation	STI	%ALcons
Bad	0.20 to 0.34	24.3 to 57
Poor	0.35 to 0.50	11.3 to 24.2
Fair	0.51 to 0.64	5.1 to 11.2
Good	0.65 to 0.86	1.6 to 5.0
Excellent	0.87 to 1.00	0.0 to 1.5

Copied from *The Audio System Designer Technical Reference* by Peter Mapp and published by Klark Teknik.

18.7.3 Architecture and Room Acoustics

18.7.3.1 Reverberation

Reverberation is the persistence of sound in a space after the original sound has been removed.

RT60 is the measure for reverberation, and it is defined as the amount of time required for the average sound energy density in a space to decrease from it's original value by 60 dB after the original sound has stopped.

The Sabine equation relates RT60 to the volume of a room with it surface area and the absorption coefficients of the materials applied to the surfaces.

As room volume increases relative to surface area and absorption coefficients, the RT60 increases.

As surface area and absorption increase relative to room volume, RT60 decreases. It is this persistence of sound that interferes with our comprehension of consonants and contributes towards degrading intelligibility.

Table 17-1. Intelligibility Comparison Chart

RT60	< 1 s	Excellent intelligibility can be achieved.
RT60	1 to 1.2 s	Excellent to good intelligibility is possible.
RT60	1.2 to 1.5 s	Good intelligibility can be achieved.
RT60	>1.5 s	Careful system design is required.
RT60	>1.7 s	Limit for good intelligibility in large spaces.
RT60	>2 s	Very directional loudspeakers are required, intelligibility can have limitations.
RT60	>2.5 s	Intelligibility will probably have limitations.
RT60	>4 s	Highly directional loudspeakers will be required to achieve acceptable intelligibility.

18.7.4 Line Arrays

Figs. 18-14 to 18-16 show the direct sound coverage of various loudspeakers in a sanctuary 100 ft × 65 ft. The chancel adds 20 ft to its length. The roof peaks at 52 ft. The room volume is roughly 250,000 ft^3. The room has plaster walls, wood ceiling, terrazzo floors, and empty wooden pews. This produces a RT60 of about 3.5 s.

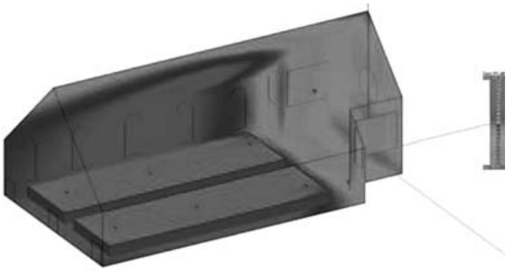


Figure 18-14. Flown large format horn array.

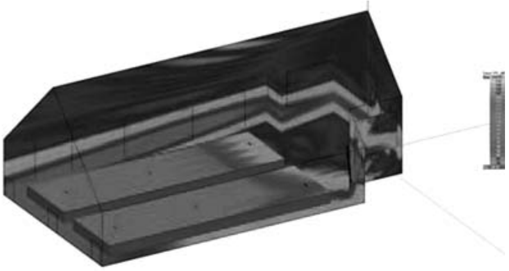


Figure 18-15. Mechanically tilted four meter column array.

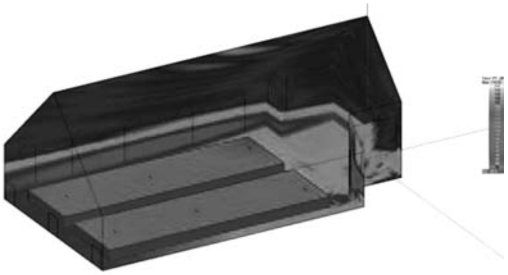


Figure 18-16. Digitally steered column array.

Notice the high SPL levels on the walls and ceiling in the flown-horn array simulation. The high frequency beaming of the mechanically tilted column array prevents good coverage of the front of the audience area. The digitally steerable column array covers only the audience area and has very little coverage on the walls and no coverage of the ceiling. Only the steered column array has acceptable (good to fair) intelligibility throughout the audience areas. Digitally steerable column arrays can offer superior coverage and they can provide improved D/R. They can provide improved intelligibility in highly reverberant spaces, plus they blend better with their surrounding architecture and are nearly invisible in use.

18.7.4.1 Digitally Steered Column Arrays

When the room size and volume are fixed and adding absorption to reduce the RT times is not an option, digitally steerable column arrays offer a new solution:

- They have the ability to be much more directional than the largest horns.
- The idea is not new; the concepts for these column arrays were described by Harry Olson in 1957. Only the implementation is new.
- The hardware required to implement these ideas is now available.
- Digital Signal Processing required is now a mature technology, very powerful and relatively inexpensive.
- Compact, highly efficient Class D amplifiers are capable of high-fidelity performance.

Line Arrays are not a new idea. Harry F. Olson did the math and described the directional characteristics of a continuous line source in his classic *Acoustical Engineering*, first published in 1940. Traditional column loudspeakers have always made use of line source directivity.

Simple line arrays (column arrays) are basically a number of drivers stacked closely together in a line, Fig. 18-17. Simple line arrays become increasingly directional in the vertical plane as the frequency increases. The spacing between drivers controls the high frequency limits. The height (length) of the line array determines the low frequency control limit. Fig. 18-18 shows the line source directivity as described by Harry Olson in 1957.

The directivity of a line array is a function of the line length and the wavelength. As the wavelength approaches the line length, the array becomes omnidirectional, Fig. 18-19. Fig. 18-20 shows the vertical dispersion pattern of a typical line array

18.7.4.2 Controlling High Frequency Beaming

Simple line arrays become increasingly directional as the frequency increase., in fact, at higher frequencies they become too directional. The vertical directivity can be made more consistent by making the array shorter as the frequency increases by using fewer drivers. One amplifier channel and one DSP channel per driver make this possible.

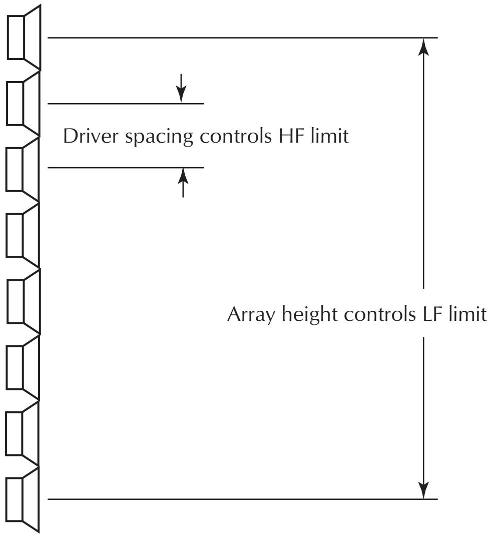

Driver spacing controls HF limit
Array height controls LF limit

Figure 18-17. Basic line array theory.

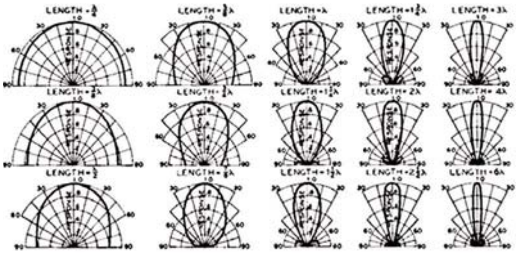


Figure 18-18. Directional characteristics of a line source as a function of the length and the wavelength. The polar graph depicts the sound pressure at a large fixed distance, as a function of angle. The sound pressure for the angle 0° is arbitrarily chosen as unity. The direction corresponding to the angle 0° is perpendicular to the line. The directional characteristics in 3D are surfaces of revolution about the line as an axis. (From Acoustical Engineering by Harry Olson.)

17.7.4.3 Beam Steering

The beam can be steered up or down by delaying the signal to adjacent drivers. DSP control also allows us to develop multiple beams from a single line array and individually steer these beams.

DSP control also allows us to move each beams acoustic center up and down the column allowing us to

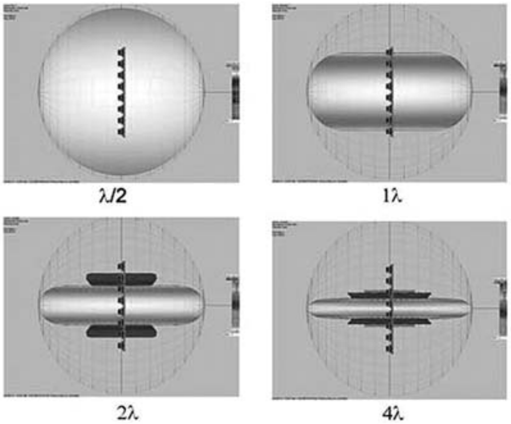


Figure 18-19. Simple line source directivity as a function of line length versus frequency.

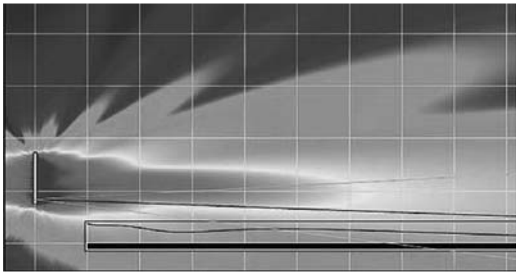


Figure 18-20. Typical line array vertical dispersion display.

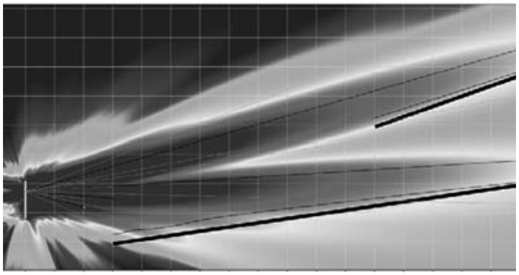


Figure 18-21. Vertical dispersion display showing multiple beam capability.

create multiple beams and also steer the beam, Figs.18-21 and 18-22.

18.7.5 DSP-Driven Vertical Arrays

18.7.5.1 Acoustical, Electronic & Mechanical Considerations

Practical examples are taken from the new Renkus-Heinz IC Series Iconyx steerable column arrays. Iconyx is a steerable column array that combines very high directivity with accurate reproduction of source material in a compact and architecturally pleasing package, Fig. 18-23.

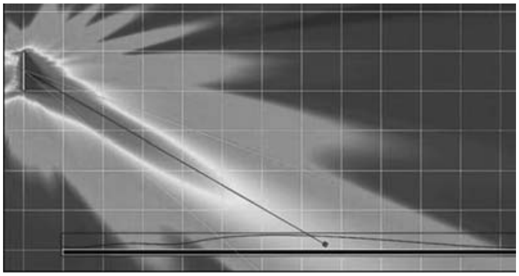


Figure 18-22. Graphic illustration of beam steering.

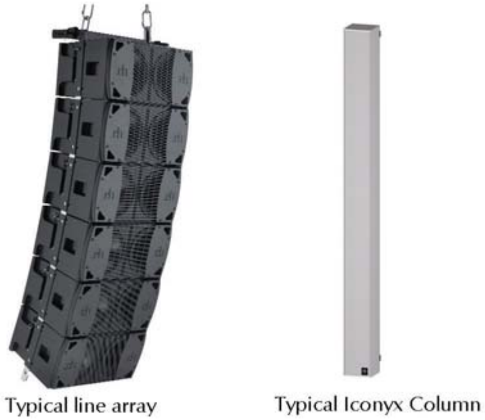


Figure 18-23. Typical line array and a typical Iconyx Column.

Like every loudspeaker system, Iconyx is designed to meet the challenges of a specific range of applications. Many of the critical design parameters are, of course, determined by the nature of these target applications. To understand the decisions that have been made during the design process we must start with the particular problems posed by the intended applications.

The function of individual driver control and DSP is to make more effective use of this phenomenon. No amount of silicon can get around the laws of acoustical physics. The acoustical properties of first-generation column loudspeakers are set by the acoustical characteristics of the transducers and the physical characteristics of the package:

1. The height of the column determines the lowest frequency at which it exerts any control over the vertical dispersion.

2. The inter-driver spacing determines the highest frequency at which the array acts as a line source rather than a collection of separate sources.
3. Horizontal dispersion is fixed and is typically set when the drivers are selected, because column loudspeakers do not have waveguides.
4. Other driver characteristics such as bandwidth, power handling and sensitivity will determine the equivalent performance characteristics of the system.

One unfortunate corollary of these characteristics is that the power response of a conventional column loudspeaker is not smooth. It will deliver much more low-frequency energy into the room and this energy will tend to have a wider vertical dispersion. This can make the critical distance even shorter because the reverberant field contains more low-frequency energy, making it harder for the listener to recognize higher-frequency sounds such as consonants or instrumental attack transients.

18.7.5.2 Point Source Interactions

17.7.5.3 Doublet Source Directivity

Doublet source cancels each other's output directly above and below, because they are spaced ½ wavelength apart in the vertical plane. In the horizontal plane, both sources sum. The overall output looks like Fig. 18-24.

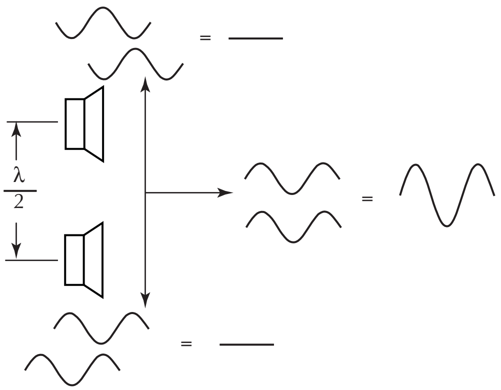


Figure 18-24. Output of a signal whose wavelength is ½ of the space between the two loudspeakers.

When two sources are ¼ wavelength apart or less, they behave almost like a single source. There is very slight narrowing in the vertical plane, Fig. 18-25.

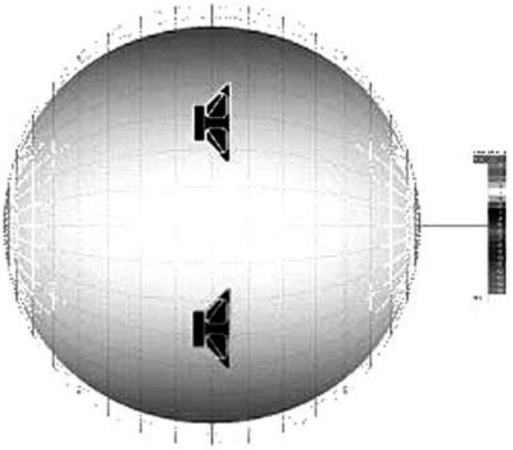


Figure 18-25. λ/4 (¼ wavelength).

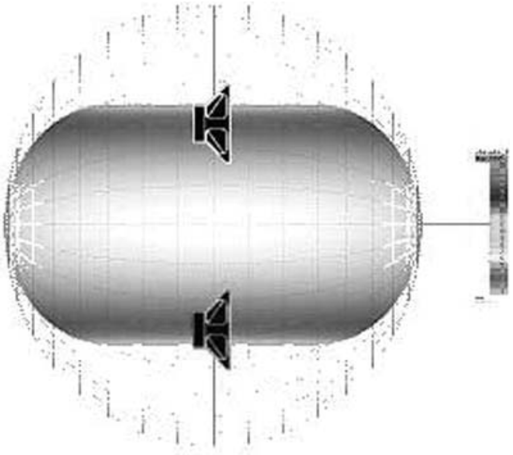


Figure 18-26. λ/2 (½ wavelength).

There is significant narrowing in the vertical plane at ½ wavelength spacing, because the waveforms cancel each other in the vertical plane, where they are 180° out of phase, Fig. 18-26.

At one wavelength spacing the two sources reinforce each other in both the vertical and horizontal directions. This creates two lobes, one vertical and the other horizontal, Fig. 18-27.

As the ratio of wavelength to inter-driver spacing increases, so do the number of lobes. With fixed drivers as used in line arrays, the ratio increases as frequency increases ($\lambda = c/f$ where f is the frequency and c is the speed of sound), Fig. 18-28.

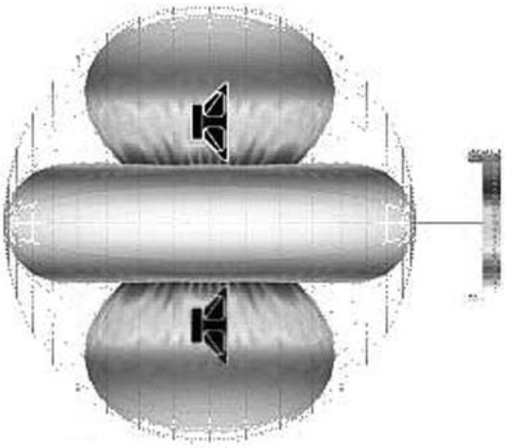


Figure 18-27. λ (1 wavelength).

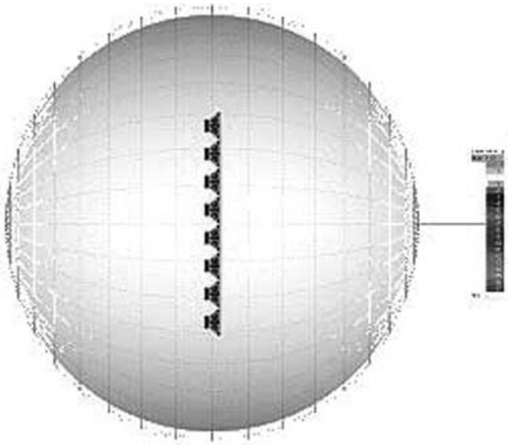


Figure 18-29. Wavelength is twice the loudspeaker height.

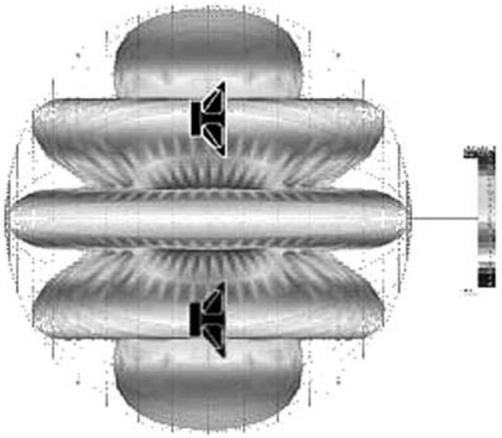


Figure 18-28. Increased wavelength to inter-driver spacing.

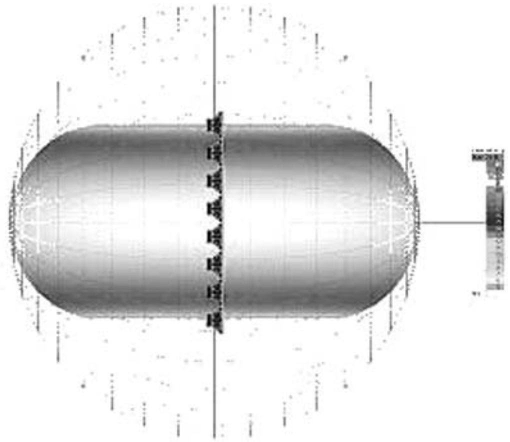


Figure 18-30. Wavelength is the loudspeaker height.

18.7.5.4 Array Height versus Wavelength (λ)

Driver-to-driver spacing sets the highest frequency at which the array operates as a line source. The total height of the array sets the lowest frequency at which it has any vertical directivity.

Figs. 18-29 though Fig. 18-32 show the effect of array height versus wavelength.

At wavelengths of twice the array height, there is no pattern control, the output is that of a single source with very high power handling, Fig. 18-29.

As the frequency rises, wavelength approaches the height of the line. At this point there is substantial control in the vertical plane, Fig. 18-30.

At higher frequencies the vertical beamwidth continues to narrow. Some side lobes appear but the energy radiated in this direction is not significant compared to the front and back lobes, Fig. 18-31.

Still further vertical narrowing, with side lobes becoming more complex and somewhat greater in energy, Fig. 18-32.

18.7.5.5 Inter-Driver Spacing versus Wavelength (λ)

The distinction between side lobes and grating lobes should to be maintained. Side lobes are adjacent to and radiate in the same direction as the primary lobe. Grating lobes are the strong summations tangential to the primary lobe. Side lobes will be present in any realizable line array, grating lobes form when the inter driver spacing becomes less than ½ wavelength. It

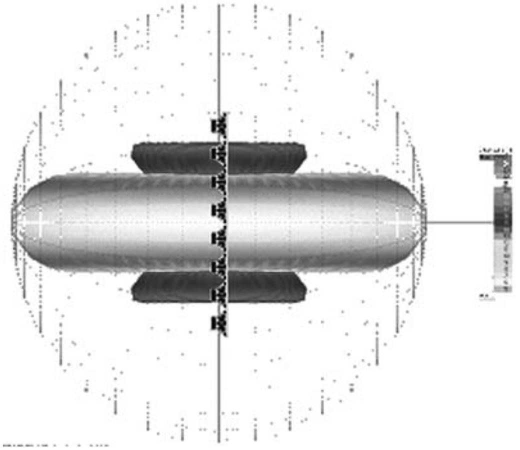


Figure 18-31. Wavelength is one half the loudspeaker height.

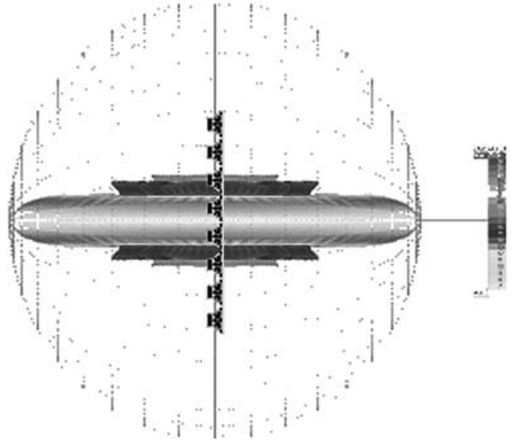


Figure 18-32. Wavelength is one fourth the loudspeaker height.

might also be good to point out that all of the graphics for this section are done using theoretical point sources.

Figs. 18-33 though Fig. 18-36 show the effect of inter-driver spacing versus wavelength.

When the drivers are spaced no more than ½ wavelength apart, the array produces a tightly directional beam with minimal side lobes, Fig. 18-33.

As the frequency rises, wavelength approaches the spacing between drivers. At this point, grating lobes become significant in the measurement. They may not be a problem, if most or all of the audience is located outside these vertical lobes, Fig. 18-34.

At still higher frequencies, lobes multiply and it becomes harder to isolate the audience from the lobes or

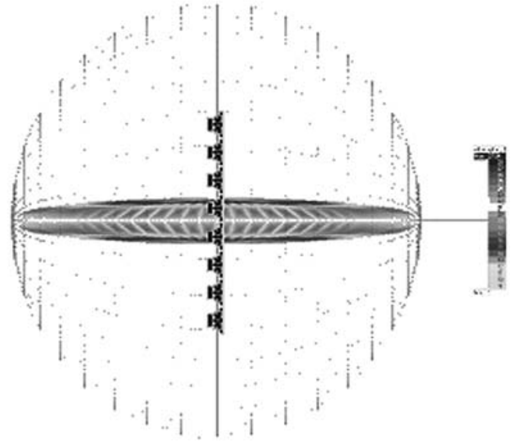


Figure 18-33. Interspacing is one half the wavelength.

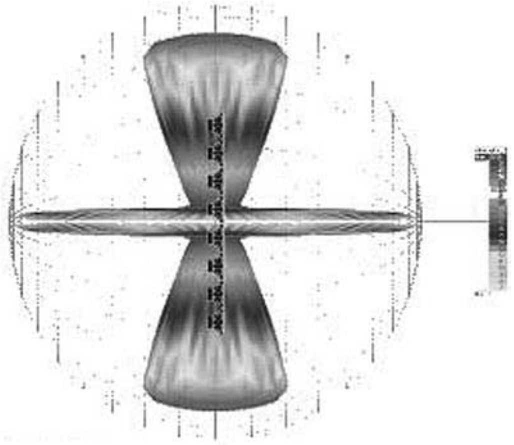


Figure 18-34. Interspacing is one times the wavelength.

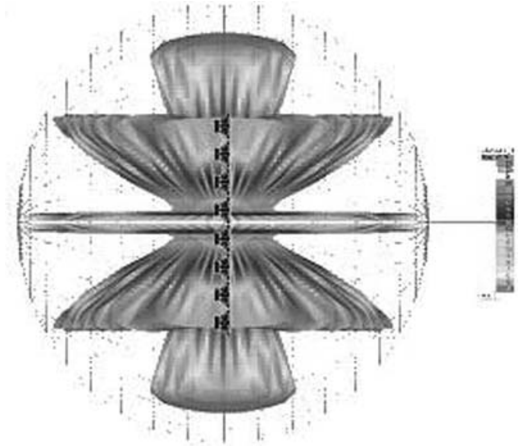


Figure 18-35. Interspacing is two times the wavelength.

As inter-driver spacing approaches four times the wavelength, the array is generating so many grating lobes of such significant energy that its output closely approximates a single point source, Fig. 18-36. We have come full circle to where the array's radiated energy is about the same as it was when array height was ½ λ. As shown in Fig. 18-32, this is the high frequency limit of line array directivity.

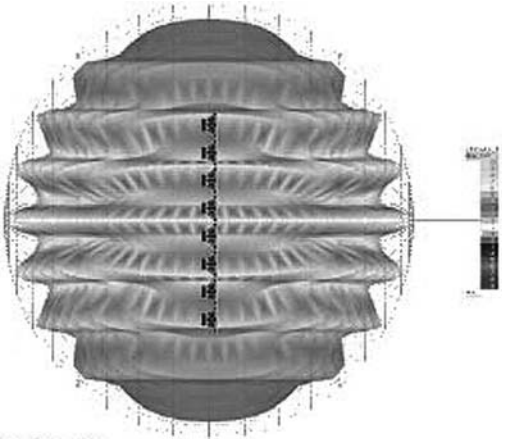


Figure 18-36. Interspacing is four times the wavelength.

As real drivers are considerably more directional than point sources at the frequencies where grating lobes are generated, the grating lobes are much lower in level than the primary lobe, Figs. 18-37 and 18-38.

18.7.6 Multichannel DSP Can Control Array Height

The upper limit of a vertical array's pattern control is always set by the inter-driver spacing. The design challenge is to minimize this dimension while optimizing frequency response and maximum output and do it without imposing excessive cost. Line arrays become increasingly directional as frequency increases, in fact, at high frequencies they are too directional to be acoustically useful. However, if we have individual DSP available for each driver, we can use it to make the array acoustically shorter as frequency increases—this will keep the vertical directivity more consistent. The technique is conceptually simple—use low-pass filters to attenuate drive level to the transducers at the top and bottom of the array, with steeper filter slopes on the extreme ends and more gradual slopes as we progress to the center. As basic as this technique is, it is practically impossible without devoting one amplifier channel and one DSP channel to each driver in the array.

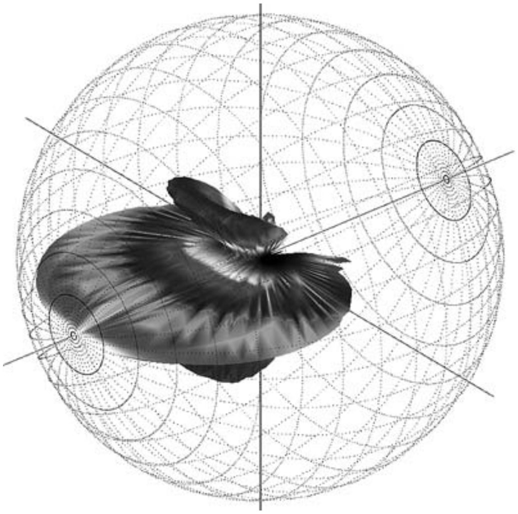


Figure 18-37. 3D view of a second generation Iconyx array at 4000 Hz.

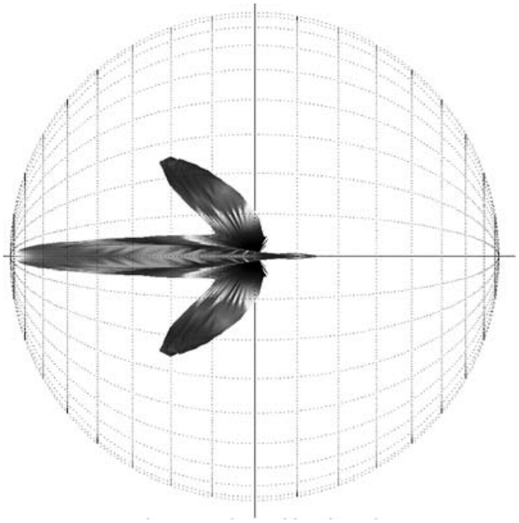


Figure 18-38. Side view of a second generation Iconyx array at 4000 Hz.

A simplified schematic shows how multichannel DSP can shorten the array as frequency increases. For clarity, only half the processing channels are shown and delays are not diagrammed, Fig. 18-39.

18.7.7 Steerable Arrays May Look Like Columns But They are not

Simple column loudspeakers provide vertical directivity, but the height of the beam changes with frequency. The overall Q of these loudspeakers is there-

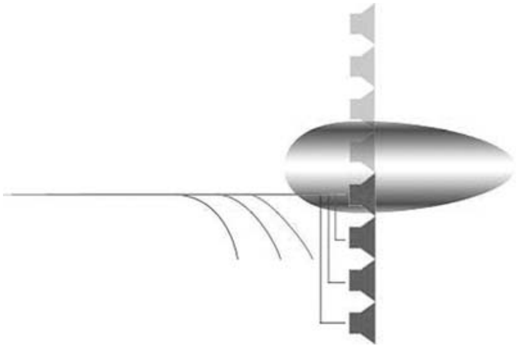


Figure 18-39. Multichannel DSP shortens the loudspeaker length.

fore lower than required. Many early designs used small-cone full range transducers, and the poor high-frequency response of these drivers certainly did nothing to enhance their reputation.

18.7.7.1 Beam-Steering: Further Proof that Everything Old is New Again

As Don Davis famously quotes Vern Knudsen, “The ancients keep stealing our ideas.” Here is another illustration from Harry F. Olson’s *Acoustical Engineering*. This one shows how digital delay, applied to a line of individual sound sources, can produce the same effect as tilting the line source. It would be long after 1957 before the cost of this relatively straightforward system became low enough for commercially viable solutions to come to market, Fig. 18-40.

18.7.7.2 DSP-Driven Arrays Solve Both Acoustical and Architectural Problems

17.7.7.3 Variable Q

DSP-driven line arrays have variable Q because we can use controlled interference to change the opening angle of the vertical beam. The Renkus Heinz IC Series can produce 5°, 10°, 15° or 20° opening angles if the array is sufficiently tall (an IC24 is the minimum required for a 5° vertical beam). This vertically narrow beam minimizes excitation of the reverberant field because very little energy is reflected off the ceiling and floor.

17.7.7.4 Consistent Q with Frequency

By controlling each driver individually with DSP and independent amplifier channels, we can use signal processing to keep directivity constant over a wide operating band. This not only minimizes the reverberant energy in the room, but delivers constant power response. The combination of variable Q, which is much higher than that of an unprocessed vertical array, with consistent Q over a relatively wide operating band, is the reason that DSP-driven Iconyx arrays give acoustical results that are so much more useful.

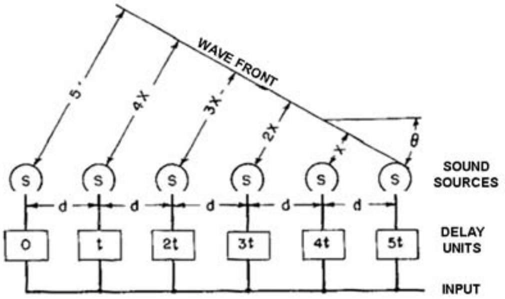


Figure 18-40. A delay system for tilting the directional characteristic of a line sound source. (From Acoustical Engineering by Harry Olson.)

17.7.7.5 Ability to Steer the Acoustic Beam Independently of the Enclosure Mounting Angle

Although beam-steering is relatively trivial from a signal-processing point of view, it is important for the architectural component of the solution. A column mounted flush to the wall can be made nearly invisible, but a down-tilted column is an intrusion on the architectural design. Any DSP-driven array can be steered. Iconyx also has the ability to change the acoustic center of the array in the vertical plane which can be very useful at times.

17.7.7.6 Design Criteria: Meeting Application Challenges

The previous figures make it clear that any line source, even with very sophisticated DSP, can control only a limited range of frequencies. However, by using full range coaxial drivers as the line source elements could make the overall sound of the system more accurate and natural without seriously compromising the benefits of beam-shaping and steering. In typical program material, most of the energy is within the range of controllable frequencies. Earlier designs radiate only slightly above and below the frequencies that are controllable. Thus much of the program source is sacrificed, without a significant increase in intelligibility.

To maximize the effectiveness of a digitally controlled line source, it's not enough to start with high quality transducers. The Renkus Heinz Iconyx loudspeaker system uses a compact multichannel amplifier with integral DSP capability. The D2 audio module has the required output, full DSP control, and the added advantage of a purely digital signal path option. When PCM data is delivered to the channel via an AES/EBU or CobraNet input, the D2 audio processor/amplifier converts it directly into PWM data that can drive the output stage.

17.7.7.7 Horizontal Directivity is Determined by the Array Elements

Vertical arrays, including Iconyx, can be steered only in the vertical plane. Horizontal coverage is fixed and is determined by the choice of array elements. The transducers used in Iconyx modules have a horizontal dispersion that is consistent over a wide operating band, varying between 140° and 150° from 100 Hz to 16 kHz.

17.7.7.8 Steering is Simple—Just Progressively Delay Drivers

If we tilt an array, we move the drivers in time as well as in space. Consider a line array of drivers that is hinged at the top and tilted downward. Tilting moves the bottom drivers further away from the listener in time as well as in space. We can produce the same acoustical effect by applying progressively longer delays to each driver as we move from top to bottom of the array.

Again, steering is not a new idea. It is different from mechanical aiming–front and rear lobes steer the same direction.

18.7.7.9 BeamWare: The Software That Controls Iconyx Linear Array Systems

A series of low-pass filters can maintain constant beamwidth over the widest possible frequency range. The ideas are simple, but for the most basic Iconyx array, the IC16, we must calculate and apply 16 sets of FIR filters, and 16 separate delay times. If we intend to take advantage of constant inter-driver spacing to move the acoustical center of the main lobe above or below the physical center of the array, we must calculate and apply a different set of filters and delays. Theoretical models are necessary, but the behavior of real transducers is more complex than the model. Each of the complex calculations underlying the Iconyx beam-shaping filters were simulated, then verified by measuring actual arrays in our robotic test and measurement facility. Fortunately, the current generation of laptop and desktop CPUs are up to the task. BeamWare takes user input in graphic form (side section of the audience area, location and mounting angle of the physical array) and provides both a simulation of the array output that can be imported into EASE v4.0 or higher, and a set of FIR filters that can be downloaded to the Iconyx system via RS422 serial control. The result is a graphical user interface that delivers precise, predictable and repeatable results in real-world acoustical environments.

This page intentionally left blank

Part 4

Electronic Audio Circuits and Equipment

This page intentionally left blank

Chapter 19

Power Supplies

by Glen Ballou

19.1 Power-Supply Terminology 669
19.2 Power Supplies 672
  19.2.1 Simple dc Power Supplies 672
  19.2.2 One-Half Wave Supplies 672
  19.2.3 Full-Wave Supplies 672
  19.2.4 Three-Phase Power Supplies 674
19.3 Filters 674
  19.3.1 Capacitor Filters 674
  19.3.2 Inductive Filters 675
  19.3.3 Combination Filters 676
    19.3.3.1 Inductance-Capacitance Filters (LC) 676
    19.3.3.2 Resistance-Capacitance Filters 676
    19.3.3.3 π Filters 676
  19.3.4 Resistance Voltage Dividers 677
19.4 Regulated Power Supplies 677
  19.4.1 Simple Regulated Supplies 679
  19.4.2 Complex Power Supplies 680
  19.4.3 Switching Regulators 682
  19.4.4 Phase-Controlled Regulated Power Supplies 683
19.5 Single IC, Power Factor Corrected, Off-Line Supply 684
19.6 Synchronous Rectification Low-Voltage Power Supplies 684
  19.6.1 Diode versus Synchronous Rectifiers 685
  19.6.2 Secondary-Side Synchronous Rectifiers 688
19.7 Converters 690
19.8 Inverters 690
19.9 Ultra Capacitor (UPS) 690
  19.9.1 Ultra Capacitor Characteristics 691
19.10 Batteries 692
  19.10.1 Temperature Effects 692
  19.10.2 Cycles versus Battery Life 692
  19.10.3 Battery Voltage 692
  19.10.4 State of Charge 692
  19.10.5 False Capacity 692
  19.10.6 Ampere-Hour Capacity 693
  19.10.7 Battery Charging 693
    19.10.7.1 Equalizing 694
    19.10.7.2 Charging Voltage versus Temperature 694
    19.10.7.3 State of Charge 694
    19.10.7.4 Internal Resistance 694

19.10.8 Lead-Acid Batteries . . . 694
19.10.9 Lead-Dioxide Batteries . . . 695
19.10.10 Absorbed Glass Mat Batteries . . . 696
  19.10.10.1 A Comparison of the Three Types of Deep Cycle Batteries . . . 696
19.10.11 LeClanche (Zinc-Carbon) Batteries . . . 697
19.10.12 Nickel-Cadmium Batteries . . . 697
19.10.13 Alkaline-Manganese Batteries . . . 698
19.10.14 Mercury Dry Cells . . . 699
References . . . 699

19.1 Power-Supply Terminology

Power Supply. A device that supplies electrical power to another unit. Power supplies obtain their prime power from the ac power line or from special power systems such as motor generators, inverters, and converters.

Rectifier. A device that passes current in only one direction. The rectifier consists of a *positive anode* and a *negative cathode*. When a positive voltage is applied to the anode of the rectifier, that voltage minus the voltage across the rectifier will appear on the cathode and current will flow. When a negative voltage is applied to the anode with respect to the cathode, the rectifier is turned off and only the rectifier leakage current will flow.

Forward Resistance. The resistance of an individual cell measured at a specified forward voltage drop or current.

Forward Voltage Drop. The internal voltage drop of a rectifier resulting from the current flow through the cell in the forward direction. The forward voltage drop is usually between 0.4 Vdc and 1.25 Vdc.

Reverse Resistance. The resistance of the rectifier measured at a specified reverse voltage or current. Reverse resistance is in megohms (MΩ).

Reverse Current. The current flow in the reverse direction, usually in microamperes (μA).

Maximum Peak Current. The highest instantaneous anode current a rectifier can safely carry recurrently in the direction of the normal current flow.

The value of the peak current is determined by the constants of the filter sections. With a choke filter input, the peak current is less than the load current. With a large capacitor filter input, the peak current may be many times the load current. The current is measured with a peak-indicating meter or oscilloscope.

Maximum Peak Inverse Voltage. The maximum instantaneous voltage that the rectifier can withstand in the direction opposite to which it is designed to pass current. Referring to Fig. 19-1, when anode A of a full-wave rectifier is positive, current flows from A to C, but not from B to C because B is negative. At the instant anode A is positive, the cathodes C of A and B are positive with respect to anode B. The voltage between the positive cathode and the negative anode B is inversely related to the voltage causing the current flow. The peak value of this voltage is limited by the resistance and nature of the path between the anode B and the cathode C. The maximum value of voltage between these points, at which there is no danger of breakdown, is termed *maximum peak inverse voltage*.

The relationship between peak inverse voltage, rms value of ac input voltage and dc output voltage depends largely on the individual characteristics of the rectifier circuit. Line surges, or any other transient or waveform distortion, may raise the actual peak voltage to a value higher than that calculated for a sine-wave voltage. The actual inverse voltage (and not the calculated value) should be such as not to exceed the rated maximum peak inverse voltage for a given rectifier. A peak-reading meter or oscilloscope is useful in determining the actual peak inverse voltage.

The peak inverse voltage is approximately 1.4 times the rms value of the anode voltage for single- phase, full-wave circuits with a sine-wave input and no capacitance at the input of the filter section. For a single half-wave circuit, with a capacitor input to the filter section, the peak inverse voltage may reach 2.8 times the rms value of the anode voltage.

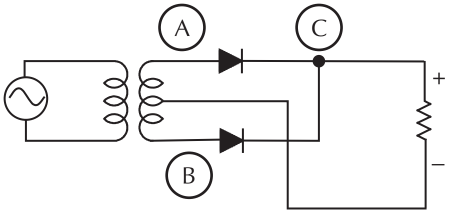


Figure 19-1. Peak inverse voltage analysis.

Ripple Voltage. The alternating component (ac) riding on the dc output voltage of a rectifier-type power supply. The frequency of the ripple voltage will depend on the line frequency and the configuration of the rectifier. The effectiveness of the filter system is a function of the load current and the values of the filter components.

The ripple factor is the measure of quality of a power supply. It is the ratio of the rms value of the ac component of the output voltage to the dc component of the output voltage or

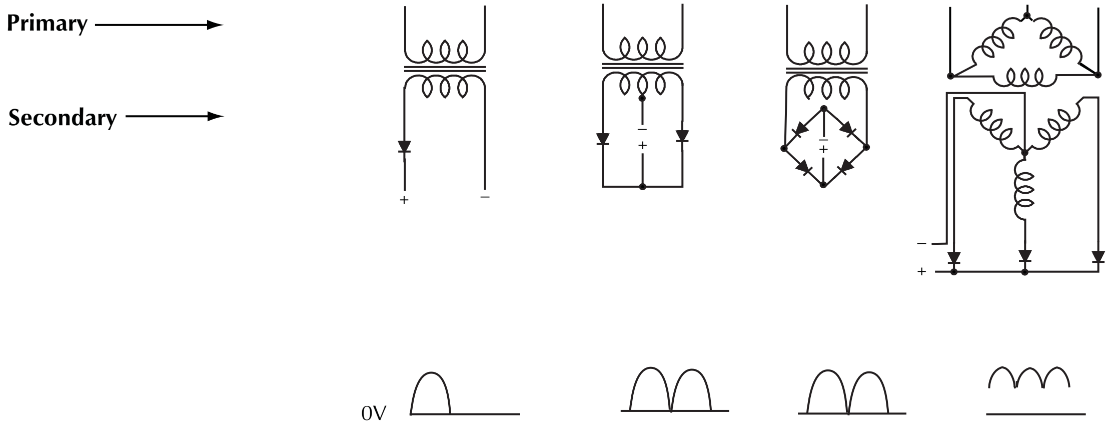
$$\textit{ripple factor} = \frac{V_{rms}}{V_{dc}} \tag{19-1}$$

where,

V_{rms} is the alternating current voltage at the output terminals,

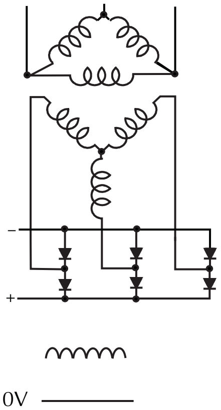
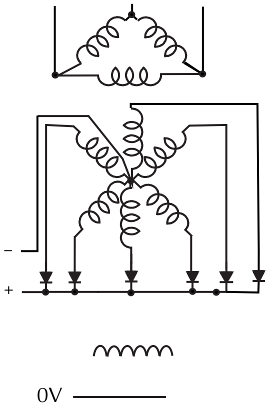
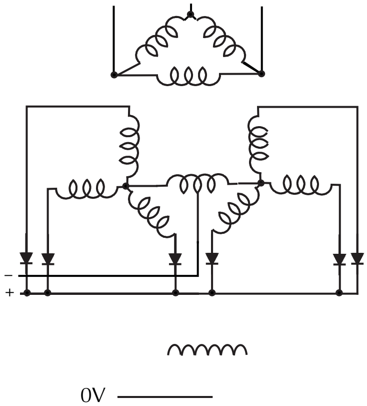
V_{dc} is the direct current output voltage at the output terminals.

Table 19-1. Rectifier Circuit Chart



Type of Circuit		Single Phase Half Wave	Single-Phase Center Tap	Single-Phase Bridge	Three-Phase Star (Wye)
Number of rectifier elements	=	1	2	4	3
Rms dc volts output	=	1.67	1.11	1.11	1.02
Peak dc volts output	=	3.14	1.57	1.57	1.21
Peak reverse volts per rectifier element	=	3.14	3.14	1.57	2.09
	=	1.41	2.82	1.41	2.45
	=	1.41	1.41	1.41	1.41
Average dc output current	=	1.00	1.00	1.00	1.00
Average dc output current per rectifier element	=	1.00	0.500	0.500	0.333
Rms current per rectifier element: Resistive load	=	1.57	0.785	0.785	0.587
Inductive load	=	–	0.707	0.707	0.578
Peak current per rectifier element: Resistive load	=	3.14	1.57	1.57	1.21
Inductive load	=	–	1.00	1.00	1.00
Ratio of peak to average current per element: Resistive load	=	3.14	3.14	3.14	3.63
Inductive load	=	–	2.00	2.00	3.00
% Ripple (rms of ripple/average output voltage)	=	121%	48%	48%	18.3%
Ripple frequency	=	1	2	2	3
		Resistive Load	Inductive Load or Large Choke Input Filter		
Transformer secondary rms volts per leg	=	2.22	1.11 (to center tap)	1.11 (total)	0.855 (to neutral)
Transformer secondary rms volts line-to-line	=	2.22	2.22	1.11	1.48
Secondary line current	=	1.57	0.707	1.00	0.578
Transformer secondary volt-amperes	=	3.49	1.57	1.11	1.48
Transformer primary rms amperes per leg	=	1.21	1.00	1.00	0.471
Transformer primary volt-amperes	=	2.69	1.11	1.11	1.21
Average of primary and secondary volt-amperes	=	3.09	1.34	1.11	1.35
Primary line current	=	1.21	1.00	1.00	0.817
Line power factor	=	–	0.900	0.900	0.826

Table 19.1 Rectifier Circuit Chart (Continued)

Three-Phase Bridge	**Six-Phase Stat (Three Phase Diametric)**	**Three-Phase Double Wye with Interphase Transformer**	
			
6	6	6	
1.00	1.00	1.00	× Average dc voltage output
1.05	1.05	1.05	× Average dc voltage output
1.05	2.09	2.42	× Average dc voltage output
2.45	2.83	2.83	× Rms secondary volts per transformer leg
1.41	1.41	1.41 (diametric)	× Rms sec. volts line-to-line
1.00	1.00	1.00	× Average dc output current
0.333	0.167	0.167	× Average dc output current
0.579	0.409	0.293	× Average dc output current
0.578	0.408	0.289	× Average dc output current
1.05	1.05	0.525	× Average dc output current
1.00	1.00	0.500	× Average dc output current
3.15	6.30	3.15	
3.00	6.00	3.00	
4.2%	4.2%	4.2%	
6	6	6	× Line frequency, *f*
Inductive Load or Large Choke Input Filter			
0.428 (to neutral)	0.740 (to neutral)	0.855 (to neutral)	× Average dc voltage output
0.740	1.48 (max)	1.71 (max-no load)	× Average dc voltage output
0.816	0.408	0.289	× Average dc output current
1.05	1.81	1.48	× Dc watts output
0.816	0.577	0.408	× Average dc output current
1.05	.28	1.05	× Dc watts output
1.05	1.55	1.26	× Dc watts output
1.41	0.817	0.707	× (Avg. load current × Sec. leg voltage)/primary line V
0.955	0.955	0.955	

Internal Output Impedance. The impedance presented to the equipment receiving the power supply voltage. In the operation of many devices, it is necessary that the internal power supply impedance be as near to zero as possible. Since most load devices consist of both passive and active elements, the current drawn from the supply consists of an ac component superimposed on the dc output of the supply. The ac component is generally of a nonsinusoidal nature. The output impedance in ohms over a wide range of frequencies is used to determine the regulation of output voltage of a power supply with respect to load variations. Power-supply output impedance (Z_o) may be defined as

$$Z_o = \frac{E_{ac}}{I_{ac}} \tag{19-2}$$

where,

E_{ac} is the sinusoidal voltage across the power supply terminals,

I_{ac} is the sinusoidal current flowing through a series loop consisting of the power supply and load equipment.

Static Line Regulation. The variation in output voltage as the input voltage is varied slowly from rated minimum to rated maximum with the load current held at the nominal value.

Dynamic Load Regulation. The variation in output when the load change is sudden. The power supply may be unable to respond instantaneously, and an additional momentary excursion in the output voltage may result, subsiding afterward to the static load regulation level. The positive and negative excursion limits are superimposed on the static line and load regulation region. The positive and negative components are not necessarily equal or symmetrical. The most stringent rating is for a change from no load to full load or from full load to no load.

Dynamic Line Regulation. The momentary additional excursion of output voltage as a result of a rapid change in input voltage.

Thermal Regulation. Variations in the output voltage over the rated operating temperature range due to ambient temperature variations influencing various components of the power supply. This is also known as *thermal drift*.

19.2 Power Supplies

19.2.1 Simple dc Power Supplies

The simplest type of dc power supply is a rectifier in series with the load. As more rectifiers are installed into the circuit, along with filters, the power supply becomes more sophisticated. The rectifier in series with the load supply will always remain simple and have poor regulation and transient response. Table 19-1 shows various power supplies and their characteristics. To determine the value of the parameter in the left column, multiply the factor shown in any of the center columns by the value in the right column.

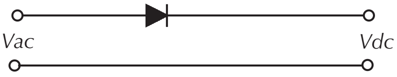
19.2.2 One-Half Wave Supplies

A one-half wave unit can be connected directly off the ac mains, Fig. 19-2A, or off the mains through a transformer, Fig. 19-2B. Since a rectifier only passes a current when the anode is more positive than the cathode, the output waveform will be one-half of a sine wave, Fig. 19-2C. The dc voltage output will be 0.45 of the ac voltage input, and the rectifier current will be the full dc current; the peak inverse voltage (*piv*) across the rectifier will be 1.414 Vac, and the ripple will be 121%. In the transformerless power supply, the 115 Vac power line is connected directly to the rectifier system. This type of power supply is dangerous to both operating personnel and to grounded equipment. Also, power supplies of this type will cause hum problems that can only be solved by the use of an isolating transformer between the line and power supply.

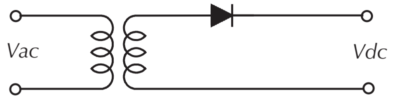
19.2.3 Full-Wave Supplies

The full-wave supply is normally used in electronic circuits because it is simple and has a good ripple factor and voltage output. A full-wave supply is always used with a transformer. Full-wave supplies may be either a single-phase center tap design or a full-wave bridge. In either case, both the positive and negative cycles are rectified and mixed to produce the dc output.

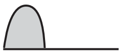
The center tap configuration, Fig. 19-2D, uses two rectifiers and a center-tapped transformer. The Vdc is approximately equal to Vac where Vac is from one side of the transformer to the center tap. Because the output is from each half wave, ripple is only 48% of the output voltage and at twice the input frequency. Each rectifier carries one-half of the load current. The piv/rectifier is 2.828 Vac.



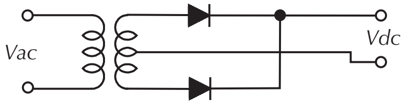
A. Half-wave transformerless power supply.



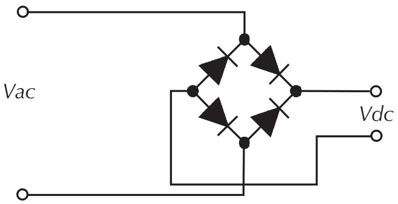
B. Half-wave transformer-isolated power supply.



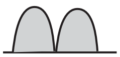
C. Output waveform of a half-wave power supply.



D. Full-wave center-tapped power supply.



E. Full-wave bridge-rectifier power supply.



F. Output waveform of a full-wave power supply.

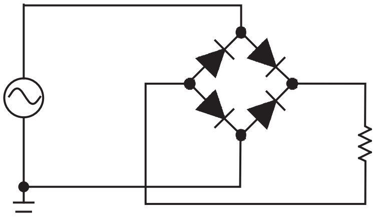
Figure 19-2. Two one-half wave and two full-wave power supplies.

A full-wave bridge rectifier supplies full-wave rectification without a center tap on the transformer. The bridge rectifier is not a true single-ended circuit, since it has no terminal common to both the input and output circuits.

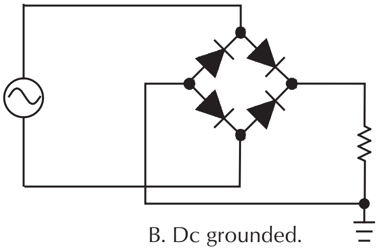
A full-wave bridge rectifier consists of four rectifier elements, as shown in Fig. 19-2E. This circuit is the most familiar and is the type most commonly employed in the electronics industry.

With the full-wave bridge circuit, the dc output voltage is equal to 0.9 of the rms value of the ac input voltage.

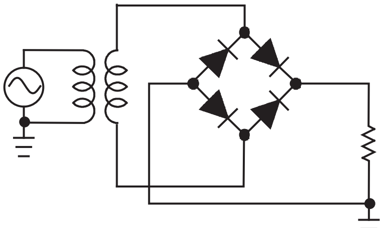
Full-wave bridge rectifier circuits may be grounded by three methods shown in Fig. 19-3A, B, C. Either the input (ac source) or output (dc load) may be grounded, but not both simultaneously. If an isolation transformer is used between the ac source and the input to the rectifier, as shown in Fig. 19-3C, both ac and dc sides may be grounded permanently. A method of grounding a bridge rectifier is shown in Fig. 19-3D where the center tap of an isolation transformer is grounded.



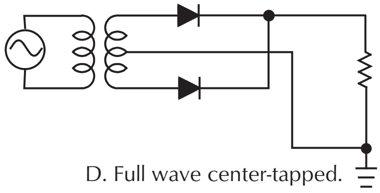
A. Ac grounded.



B. Dc grounded.



C. Ac and/or dc grounded using an isolation transformer.



D. Full wave center-tapped.

Figure 19-3. Methods of grounding a power supply.

When designing rectifier circuits, dc load current, dc load voltage, peak inverse voltage, maximum ambient temperature, cooling requirements, and overload current must be analyzed. For example, assume a full-wave rectifier using silicon rectifiers is to be designed as in Fig. 19-3D and the dc load voltage Vdc under load is 25 V at 1 A.

Using Table 19-1, determine the current per rectifier using the equation

$$\begin{aligned} I_{rect} &= 0.5 \times I_{dc} \\ &= 0.5 \times 1 \\ &= 0.5A \end{aligned} \qquad (19\text{-}3)$$

where,
I_{rect} is the current per rectifier,
0.5 is a constant from Table 19-1,
I_{dc} is the rectified ac current, which is the dc current.

This is the current each rectifier must carry. Next, the ac voltage required from the transformer is determined by the equation

$$\begin{aligned} V_{ac} &= 1.11 \times V_{dc} \\ &= 1.11 \times 25 \\ &= 27.75\ V_{rms} \end{aligned} \qquad (19\text{-}4)$$

where,
V_{ac} is the transformer voltage,
1.11 is a constant from Table 19-1.

This is the voltage as measured from each side of the transformer center tap; the total voltage across the secondary is 55.50 Vrms.

The peak inverse voltage is

$$\begin{aligned} piv &= 2.82 \times V_{ac} \\ &= 2.82 \times 27.75 \\ &= 78.4\ V_{rms} \end{aligned} \qquad (19\text{-}5)$$

where,
V_{ac} is the secondary ac voltage per leg,
2.82 is found in Table 19-1.

If a rectifier with the required *piv* rating is not available, two or more may be connected in series to obtain the desired *piv* rating. Unequal values of *piv* ratings may be used, provided the lowest rating is greater than half of the total piv rating needed.

Parallel operation of rectifiers can be used obtain higher current ratings. However, because of a possible imbalance between the units due to the forward voltage drop and effective series resistance, one unit may carry more current than the other and could conceivably fail. To prevent this, small equal value resistors must be connected in series with each individual rectifier to balance the load currents, as shown in Fig. 19-4.

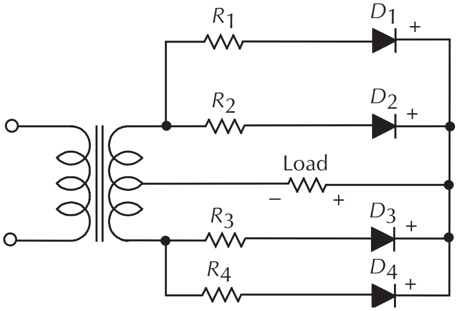


Figure 19-4. Small resistors connected in series with each rectifier to balance the current through each unit of parallel-connected rectifiers.

19.2.4 Three-Phase Power Supplies

Three-phase power supplies are common in the industry but are seldom used to power audio circuits directly. The are used as the input power to power an entire system—for instance, a portable high-power outdoor rock system. To see the characteristics of three-phase supplies, see Table 19-1.

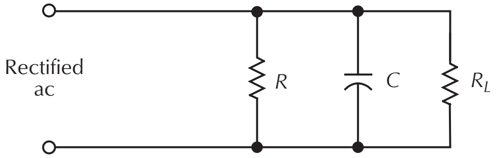
19.3 Filters

A *power-supply filter* is a series of resistors, capacitors, and/or inductors connected either passively or actively to reduce the ac or ripple component of the dc power supply.

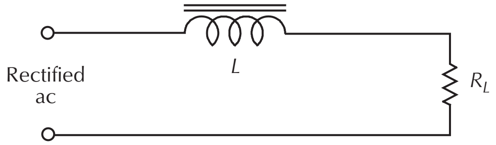
19.3.1 Capacitor Filters

A *capacitor filter* employs a capacitor at its input, as shown in Fig. 19-5A. Power supplies with an input capacitor filter have a higher output voltage than one without a capacitor because the peak value of the rectifier output voltage appears across the input filter. As the rectified ac pulses from the rectifier are applied across capacitor *C*, the voltage across the capacitor rises nearly as fast as the pulse. As the rectifier output drops, the voltage across the capacitor does not fall to zero but gradually diminishes until another pulse from the rectifier is applied to it. It again charges to the peak voltage. The capacitor may be considered a storage tank, storing up energy to the load between pulses. In a half-wave circuit, this action occurs 60 times per second, and in a full-wave circuit, it occurs 120 times per second.

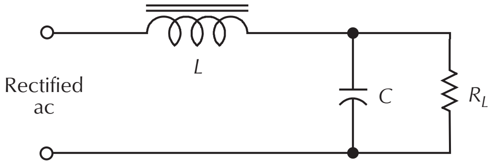
For a single-phase circuit with a sine-wave input and without a filter, the peak inverse voltage at the rectifier is 1.414 times the rms value of the voltage applied to the rectifier. With a capacitor input to the filter, the peak inverse voltage may reach 2.8 times the rms value of the applied voltage. This data may be obtained by referring to Table 19-1.



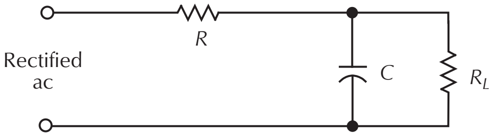
A. Capacitor filter.



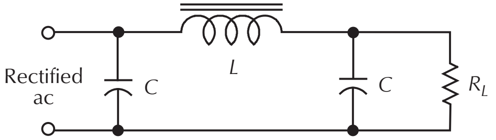
B. Choke input filter.



C. Inductance-capacitance or L filter.



D. Resistance-capacitance filter.



E. π filter using inductance and capacitance.

Figure 19-5. Capacitive, inductive, and π filters.

When a dc voltmeter is connected across the unfiltered output of a rectifier, it will read the average voltage. As an example, assume a dc voltmeter is connected across a half-wave rectifier. Because of the inertia of the meter pointer movement, the meter can not respond to the rapidly changing pulses of the half-wave rectified current but acts as a mechanical integrator. The pointer will be displaced an amount proportional to the time average of the applied voltage waveform.

The average voltage (V_{av}), as read by the dc voltmeter, is

$$V_{av} = \frac{V_p}{\pi} \tag{19-6}$$

where,
V_p is the peak voltage,
π is 3.1416….

The ripple factor is

$$\gamma = \frac{I_{dc}}{4\pi\sqrt{3}fCV_{dc}} \tag{19-7}$$

where,
I_{dc} is the output dc current,
f is the ripple frequency,
C is the filter capacitor in farads,
R_L is the load resistance in ohms.

Capacitor filters operate best with large filter capacitors and high-resistance loads. As the load resistance is lowered, the ripple increases and regulation decreases.

Filtering efficiency is reduced, and the internal leakage is increased when the capacitor's power factor increases. Electrolytic capacitors should be removed when their power factor reaches an excessive value. In an ideal capacitor, the current would lead the voltage by 90°. Capacitors are never ideal because a small amount of leakage current always exists through the dielectric. Also, a certain amount of power is dissipated by the dielectric, the leads, and their connections. All this adds up to power loss. This power loss is termed *phase difference* and is expressed in terms of power factor (*PF*). The smaller the power factor value, the more effective the capacitor. Since most service capacitor analyzers indicate these losses directly in terms of power factor, capacitors with large power factors may be readily identified. Generally speaking, when an electrolytic capacitor reaches a power factor of 15%, it should be replaced. The filtering efficiency for different values of power factor can be read directly from Table 19-2.

Table 19-2. Filtering Efficiency versus %Power Factor

Filtering Efficiency	% PF	Filtering Efficiency	% PF
100	0.000	35	0.935
90	0.436	30	0.955
80	0.600	25	0.968
70	0.715	20	0.980
60	0.800	15	0.989
50	0.857	10	0.995
45	0.895	5	0.999
40	0.915		

19.3.2 Inductive Filters

An *inductive filter* employs a choke rather than a capacitor at the input of the filter, as shown in Fig. 19-5B. Although the output voltage from this type of filter is lower, the voltage regulation is better.

A choke filter operates best with maximum current flow. It has no effect on a circuit when no current is flowing. The critical inductance is the inductance required to assure that current flows to the load at all times. An inductor filter depends on the property of an inductor to oppose any change of current.

To assure that current flows continuously, the $\sqrt{2}I_{rms}$ peak current of the ac component of the current must not exceed the direct current $I_{dc} = I_{dc}/R_L$. Therefore,

$$X_L \geq \frac{\sqrt{2}}{3R_L} \tag{19-8}$$

and

$$L_C = \frac{R_L}{3 \times 2\pi f} \tag{19-9}$$

where,
L_C is the critical inductance,
R_L is the load resistance.

Filter chokes should be selected for the lowest possible dc resistance commensurate with the value of inductance.

The ripple factor (γ) for an inductive filter is

$$\gamma = \frac{R_L + R_C}{3\sqrt{2} \times 2\pi f} \tag{19-10}$$

where,
R_L is the load resistance in ohms,
R_C is the choke resistance in ohms,
f is the ripple frequency.

19.3.3 Combination Filters

Combination filters use a combination of resistors, capacitors, and inductors to improve the filtering. The simplest is a resistor-capacitor filter and the more complicated is a series of inductance-capacitor (*LC*) circuits.

19.3.3.1 Inductance-Capacitance Filters (LC)

Inductance- capacitance filters, sometimes called *L filters*, use an inductor as an input filter and a capacitor as the second stage of the filter, Fig. 19-5C. LC filters operate well under varying load conditions.

The inductive reactance of the choke in an LC filter section tends to oppose any change in the current flowing through the winding, creating a smoothing action on the pulsating current of the rectifier. The capacitor stores and releases electrical energy, also smoothing out the ripple voltage, resulting in a fairly smooth output current.

The ripple factor for an LC filter is

$$\begin{aligned}\gamma &= \frac{\sqrt{2}X_C}{3X_L} \\ &= \frac{\sqrt{2}}{3 \times 2\pi fC \times 2\pi fL} \\ &= \frac{0.01}{f^2 CL}\end{aligned} \tag{19-11}$$

where,
X_C is the capacitance reactance in ohms,
X_L is the inductive reactance in ohms,
f is the frequency of ripple,
C is the capacitance in farads,
L is the inductance in henrys.

When multiple LC filters are connected together, the ripple factor is

$$\begin{aligned}\gamma &= \frac{\frac{\sqrt{2}}{3}}{(16\pi^2 f^2 LC)^n} \\ &= \frac{0.47}{(157.9 f^2 LC)^n}\end{aligned} \tag{19-12}$$

where,
L is the inductance in henrys,
f is the ripple frequency,
C is the capacitance in farads,
n is the number of sections.

19.3.3.2 Resistance-Capacitance Filters

Resistance-capacitance filters, **RC,** Fig. 19-5D, employ a resistor and capacitor rather than an inductor and capacitor. The advantages of such a filter are its low cost, light weight, and the reduction of magnetic fields. The disadvantage of such a filter is that the series resistance induces a voltage drop that varies with current and could be detrimental to the circuit operation. An RC filter system is generally used only where the current demands are low. RC filters are not as efficient as the LC type, and they may require two or more sections to provide sufficient filtering.

19.3.3.3 π Filters

A π (pi) *filter* has a capacitor input followed by an LC section filter, Fig. 19-5E. π filters have a smooth output and poor regulation. They are often used where the transformer voltage is not high enough and low ripple is required. By using the input capacitor, the dc voltage is boosted to the peak voltage. The ripple factor for a π filter is

$$\gamma = \sqrt{2}\,\frac{X_{C1}X_{C2}}{R_L X_{L1}} \tag{19-13}$$

where,

X_{C1} is the capacitive reactance of the first capacitor,
X_{C2} is the capacitive reactance of the second capacitor,
R_L is the load resistance,
X_{L1} is the inductive reactance of the choke.

When the choke is replaced with a resistor, the ripple factor becomes

$$\gamma = \sqrt{2}\,\frac{X_{C1}X_{C2}}{R_L R} \tag{19-14}$$

where,

R is the filter resistor.

19.3.4 Resistance Voltage Dividers

A *resistance voltage divider* is shown in Fig. 19-6. In this system of voltage division, the resistors are connected in series with the particular load they feed. The resistors are calculated by means of Ohm's law

$$R = \frac{V}{I}. \tag{19-15}$$

The wattage is computed by

$$\begin{aligned} P &= \frac{V^2}{R} \\ &= I^2 R. \end{aligned} \tag{19-16}$$

Generally, when a series-resistance voltage divider is used, a separate bleeder resistor is also used to secure better regulation. Each section should have a separate bypass capacitor of 10 μF or more to ground. The bypass capacitors stabilize and improve the filtering and decouple the various levels. This is particularly true for the series-type voltage divider.

There are two common types of voltage dividers, the *shunt* and the *series* types. The shunt type shown in Fig. 19-6 is designed to supply three different voltages to external devices. The upper circuit supplies load L_1, the second circuit supplies L_2, and the third circuit supplies L_3. All circuits are common to ground.

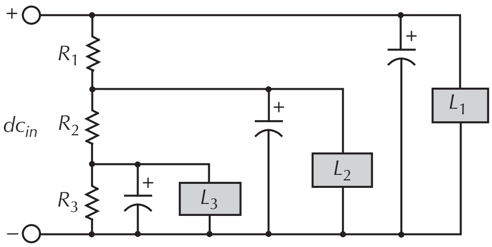


Figure 19-6. Shunt-type voltage-divider system showing the current flow in the various branches.

The total current required is the total current of the three external circuits, or $I_{L1} + I_{L2} + I_{L3}$, plus an additional current called the *bleeder current*. This bleeder current flows only through the resistors and not through the external circuits. It is generally 10% of the total current.

Resistor R_1 is calculated first, because only bleeder current flows through this resistor,

$$R_1 = \frac{V}{I} \tag{19-17}$$

where,

V is the L_1 voltage also across R_1,
I is the bleeder current.

The voltage at the top of R_2 is the L_2 voltage to ground. Subtracting the voltage drop across R_1 results in a voltage across R_2. The current through R_2 is the current of load L_1 plus the bleeder current

$$R_2 = \frac{V_{L_2} - V_{L_1}}{I_{R_1} + I_{L_1}}. \tag{19-18}$$

Resistor R_3 has the current of loads L_1 and L_2 plus the bleeder current flowing through it or

$$R_3 = \frac{V_{L_3} - V_{L_2}}{I_{R_1} + I_{L_1} + I_{L_2}}. \tag{19-19}$$

The current of load L_3 does not flow through any part of the voltage-divider system; therefore, it requires no further considerations.

19.4 Regulated Power Supplies

A *regulated power supply* holds the output constant with variations in load, current, or input voltage. Regulated supplies may be simple shunt or series regulators with 1–3% regulation or high gain supplies with 0.001% regulation and 0.001% ripple.

Power supplies may be connected in parallel, but to protect the supplies, diodes are connected in the positive lead of each power supply. When the diode is in its normal conducting mode, it must be capable of withstanding the short-circuit current of its regulator. The piv rating of the diode must be equal to or greater than the maximum open-circuit potential of the highest-rated power supply.

Regulated power supplies can also be connected in series if certain precautions are observed. The isolation voltage rating of the individual power supplies must not be exceeded, and the power supplies must be protected against reverse potential. Diodes are connected in the nonconducting direction across the output of each supply unit. These diodes will start to conduct the instant a reverse potential appears, providing a path for short-circuit current. If possible, the regulating circuit for one supply should be connected as a master and the other as slaves. The voltages of the supplies do not have to be the same.

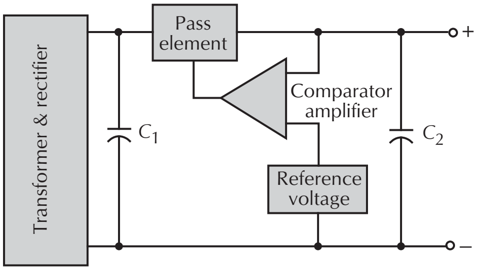
All regulated supplies have a reference element and a control element. The amount of electronics between the two elements determines the quality and regulation of the supply, Fig. 19-7.

The reference element is the unit that forms the foundation of all voltage regulators. The output of the regulated power supply is equal to or a multiple of the reference. Any variation in the reference voltage will cause the output voltage to vary; therefore, the reference voltage must be maintained as stable as possible.

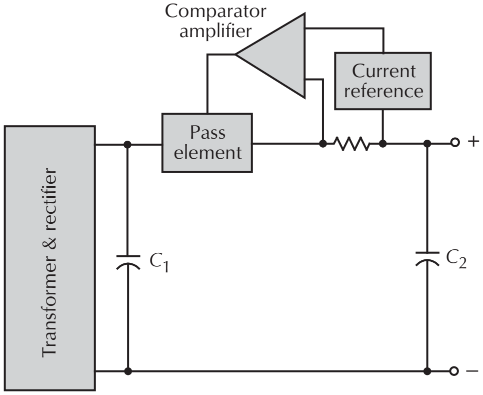
The control element is that unit that maintains the output voltage constant. The regulator type is named after the control element—namely, series, shunt, or switching, Fig. 19-7A, B, C. The control element is an electronic variable resistor that drops voltage either in series with the load or across the load. Control element configurations are shown in Fig. 19-8.

All regulated supplies draw standby current, which is the current drawn by the power supply with no output load. The input voltage to regulated supplies is filtered dc. The smoother the input voltage is, the smoother the output will be. The capacitor C_1 shown in Fig. 19-7A is used to smooth the output or reduce ripple.

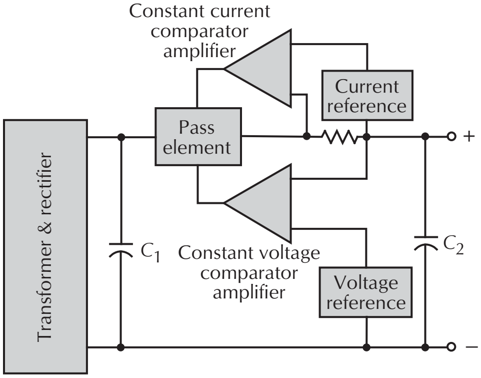
The comparison amplifier constantly monitors the output, reducing ripple because the reference voltage is smooth dc and the output ripple voltage appears to the comparator like a varying load. The regulator or pass transistor attempts to follow it, reducing ripple.



A. Block diagram for a constant voltage regulated power supply.



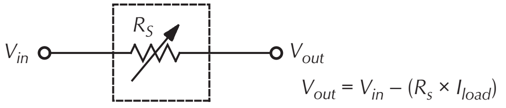
B. Block diagram for a constant current regulated power supply.



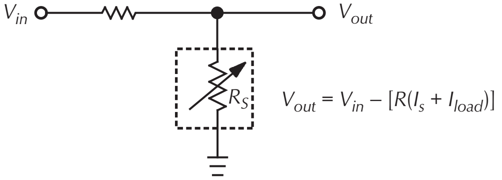
C. Constant voltage constant current regulated power supply with automatic crossover

Figure 19-7. Regulated power supplies.

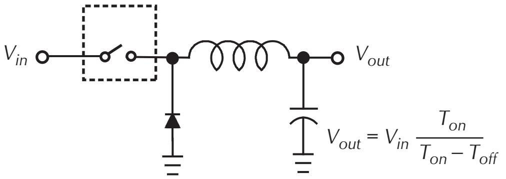
A *constant-voltage regulated power supply* is designed to keep its output voltage constant, regardless of the changes in load current, line voltage, or temperature. For a change in the load resistance, the output voltage remains constant to a first approximation, while the output current changes by whatever amount is



A. Series regulator.



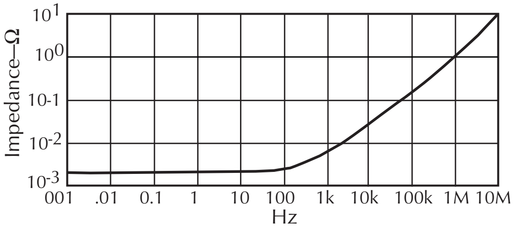
B. Shunt regulator.



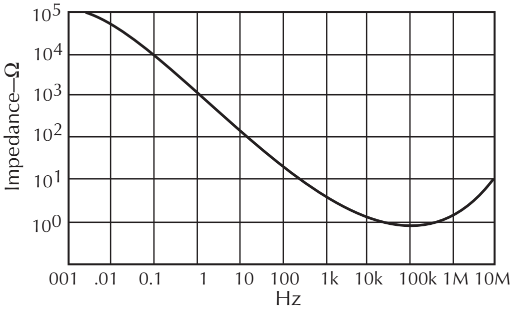
C. Switching regulator.

Figure 19-8. Examples of various control elements.

necessary to accomplish this, Fig. 19-7A. Its impedance curve is shown in Fig. 19-9A.



A. Constant voltage power supply.



B. Constant current power supply.

Figure 19-9. Typical internal impedance characteristics for a regulated power supply.

An ideal constant-voltage power supply would have zero impedance. For well-designed voltage-regulated power supplies, the internal output impedance will range from 0.001–3 Ω for frequencies from dc–1 MHz. The actual impedance is a function of the load and the type of equipment being fed by the supply.

A *constant-current regulated power supply* is designed to keep its output current constant, regardless of the changes in load impedance, line voltage, or temperature. For a change in the load resistance, the output current remains constant to a first approximation, although the output voltage changes by whatever amount is necessary to accomplish this, Fig. 19-7B. Its impedance characteristics are given in Fig. 19-9B.

A constant-current supply would have infinite impedance at all frequencies. However, these ideals are not achieved. Therefore, a practical power supply has a very low impedance at the lower frequencies, and the impedance rises with frequency. The constant-current supply has a rather high impedance at the lower frequencies and decreases at the higher frequencies.

A *constant-voltage, constant-current regulated power supply,* Fig. 19-7C, acts as a constant-voltage source for comparatively large values of load resistance and as a constant-current source for comparatively small values of load resistance. An automatic crossover (or transition) between these two modes of operation occurs at a critical or crossover value of load resistance (R_C) where

$$R_C = \frac{V_s}{I_s} \tag{19-20}$$

where,

V_s is the voltage-control setting,

I_s is the current-control setting.

19.4.1 Simple Regulated Supplies

A simple supply consists of only the control element and the reference element. The *solid-state zener diode* has almost replaced the gaseous tube reference element because it is smaller and has better regulation, wide voltage range, and wide power range. Referring to the basic design in Fig. 19-10A, the zener diode is connected in series with the limiting resistor R_1 and in parallel with the output. As a rule, the zener diode current I_Z is chosen for a value of 10% of the load current I_L. The value of the series resistance R_1 can be calculated using the equation

$$R_1 = \frac{V_s - V_{L_1 out}}{I_L + I_Z} \tag{19-21}$$

where,

V_s is the voltage source,

V_{out} is the output voltage,

I_L is the load current,
I_Z is the zener current, (normally 10% of I_L.)

The power dissipated in R_1 is I^2R. The dissipation is only for a condition where the load current remains constant at its design current. If the load current is completely removed, the current through the diode increases to the design load current plus the design zener current.

Two additional voltage-regulating circuits are shown in Fig. 19-10B and C. Zener diodes can be connected in series across the output of a dc supply, provided the power-handling capabilities and the current-operating ranges are similar.

A cascade shunt regulator is given in Fig. 19-10D. The zener diode controls the base potential of transistor Q_1, which functions as an emitter follower and circuit amplifier. This circuit is used where large current variations are encountered.

If only a small voltage drop is required, i.e., 5–6 V, the configuration in Fig. 19-10E might be employed. In this instance, the entire load current plus the current through R_1 must flow through the diode, and it could be easily damaged.

A current-regulator circuit is shown in Fig. 19-10F. The load current remains essentially constant until R_L increases to where the average voltage drop across R_L is as large as the voltage drop across R_3.

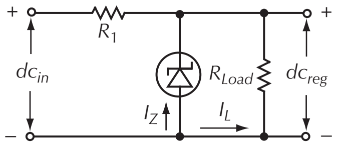
19.4.2 Complex Power Supplies

Complex supplies include a pass element, a sampling element, and a comparator element, and they may include a preregulator, current limiting, undervoltage and overvoltage protection, and remote sensing.

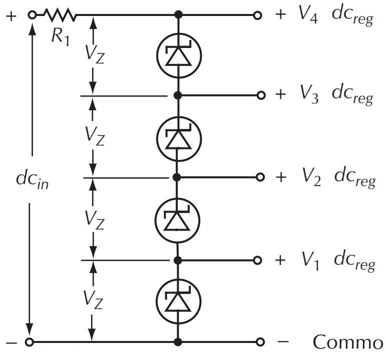
Pass Elements. A transistor or group of transistors connected in parallel and placed in series with the output of a regulated power supply to control the flow of the output current. A *pass element* is another name for control element.

Reference Elements. The unit that forms the foundation of all voltage regulators. The output of the regulated power supply is equal to or a multiple of the reference. Any variation in the reference voltage will cause the output voltage to vary; therefore, the reference voltage must be maintained as stable as possible.

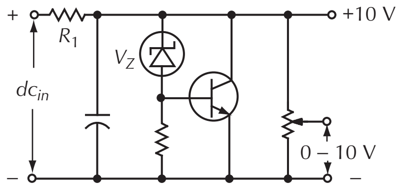
Sampling Elements. The device that monitors the output voltage and translates it into a level comparable to the reference voltage. The variations in the sampling voltage versus the reference voltage is the error voltage



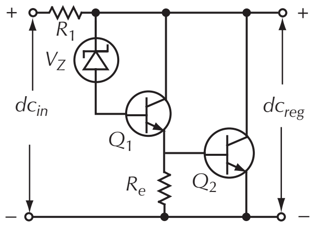
A. Simple zener voltage regulator circuit.



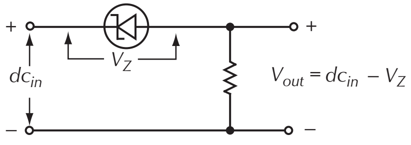
B. Zener diodes connected in series for regulation and voltage division circuits.



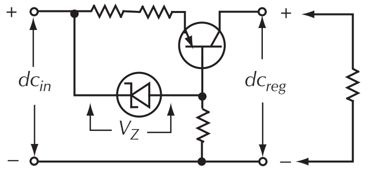
C. Zener diode regulator circuit where voltages lower than the zener diode are desired.



D. Cascade shunt zener-diode voltage-regulator circuit.



E. Series connection for a zener diode when only a small voltage drop is required.



F. Current-regulator using a transistor and zener diode.

Figure 19-10. Various regulator circuits using zener diodes.

that ultimately controls the regulator output.

Comparator Elements. Compares the feedback voltage from the sampling element with the reference and provides gain for the detected error level. This signal controls the control circuit.

Preregulator. Monitors the voltage across the series regulator and adjusts the input V_{in} to maintain the regulator voltage at approximately 3 volts. This regulator voltage is held relatively constant regardless of input or output conditions. This reduces the power dissipated and the number of transistors in the series regulator, Fig. 19-11.

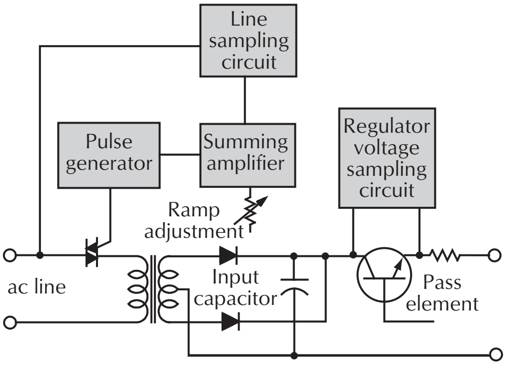


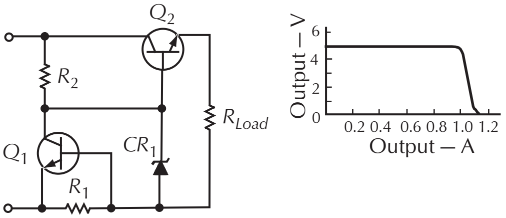
Figure 19-11. Simplified block diagram of a preregulated power supply.

Current Limiting. A method used to protect the pass transistor by limiting the current within the safe operating range. The simplest current-limiting device is a resistor in series with the load. This, however, affects regulation by the IR drop across the resistor.

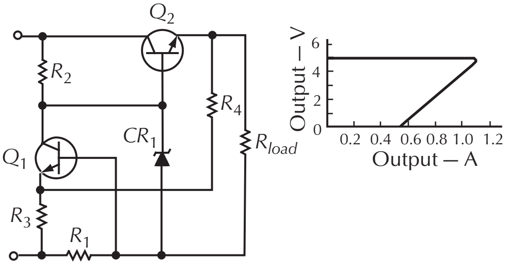
To overcome this, constant current limiting is used. With constant current limiting, the voltage drop across the series resistor is sampled. The output voltage remains constant up to a predetermined current at which time the voltage decreases to limit the output current.

A third current limiting is foldback current limiting in which the load current actually decreases as the load continues to increase beyond I_{max}. This is usually only used in high current supplies.

The conventional current-limiting power supply of Fig. 19-12A is protected from instantaneous short circuits but long duration shorts can overheat Q_2, leading to its eventual failure. In Fig. 19-12B this circuit is modified to produce foldback by adding two voltage feedback resistors, R_3 and R_4. The control transistor Q_1 emitter voltage depends on the power supply output voltage as sampled by the R_3, R_4 voltage divider. If R_1 senses a current overload, the drop across it decreases the output voltage and lowers the emitter voltage of Q_1. Then Q_1 turns on at reduced current through R_1, which limits current flow through Q_2, as shown in the current-foldback characteristic of Fig. 19-12B. The foldback ratio can be adjusted by changing R_3, R_4, or R_1, or all three.



A. Conventional current limiting circuit.



B. Addition of feedback resistors to generate a current foldback output.

Figure 19-12. Current limiting circuits.

Overvoltage Protection. Protects the load from overvoltage. This may be accomplished internally or as an add-on to the power supply. A crowbar circuit is a typical overvoltage protector.

The circuit monitors the output voltage of a power supply and instantaneously throws a short circuit across the output terminals when a preset voltage is reached. This is generally accomplished by the use of a silicon controlled rectifier (SCR) connected across the output terminals of the supply unit.

Remote Sensing. Adding two extra wires between the supply and the load produces a *remote sensing* circuit that permits the supply to achieve its optimum regulation at the load terminals, rather than at the power supply output terminals. In this manner, the circuit compensates for the IR drop in the line from the power supply to the equipment receiving its voltage. The sensing lines are high impedance and have almost no current flowing. Therefore, the voltage drop is negligible.

The wire size and voltage drop for regulated power supplies can be determined by Ohm's Law or with the use of the nomograph in Tables 14-2 and 14-3. Since regulated power supplies are designed to control the output at the power supply output terminals, the conductors used for the supply line must be considered as a part of the power supply load.

19.4.3 Switching Regulators

In a switching regulator, the pass transistor operates in an on-off mode, increasing efficiency and reducing heat. The simple switching regulator shown in Fig. 19-13 incorporates a pulse generator circuit that pulses on the pass transistor as the output voltage decreases. As the output voltage increases, the comparator circuit reduces the pulse generator, reducing the on time of the pass transistor and, therefore, reducing the average output voltage. Since the output voltage is a series of pulses, a filter is required to smooth the dc output. An inductance-capacitance filter is commonly used. Switching regulators normally operate at 20 kHz or higher and have the following advantages:

- Switching regulators are on-off devices, so they avoid the higher power dissipation associated with the rheostat like action of a series regulator. Transistors dissipate very little power when either saturated (on) or nonconducting (off); most of the power losses occur elsewhere in the supply. Efficiency to 85% is typical for switching supplies, as compared to 30–45% for linear supplies. Less wasted power means switching supplies run cooler, cost less to operate, and have smaller regulator heat sinks.
- Size and weight reductions for switching supplies are achieved because of their high switching rate. Typically, a switching supply is less than one-third the size and weight of a comparable series-regulated supply.
- Switching supplies can operate under low ac input voltage (brownout) conditions and sustain a relatively long carryover (or holdup) of its output if input power is lost momentarily because more energy is stored in its input filter capacitors. In a switching supply, the input ac voltage is rectified directly, and the filter capacitor charges to the ac voltage peaks. The ac voltage input of the standard linear supply is stepped down through a power transformer and then rectified, resulting in a lower voltage across its filter capacitor. Since the energy stored in a capacitor is proportional to CV^2 and V is higher in switching supplies, their storage capability (and thus their holdup time) is better.

Switching disadvantages include the following:

- A switching supply transient recovery time (dynamic load regulation) is slower than that of a series-regulated supply. In a linear supply, recovery time is limited only by the speeds of the semiconductors used in the series regulator and control circuitry. In a switching supply, recovery is limited mainly by the inductance in the output filter.
- Electromagnetic interference (emi) is a natural byproduct of the on-off switching. This interference can be conducted to the load (resulting in higher output ripple and noise), it can be conducted back into the ac line, and it can be radiated into the surrounding atmosphere.

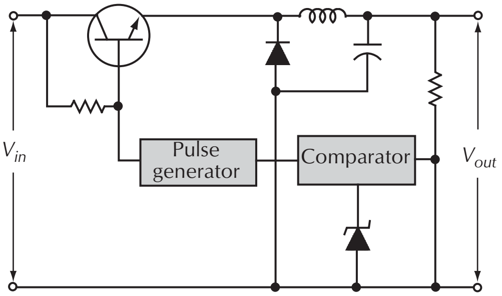


Figure 19-13. Basic switching regulator.

High-Power Regulated Supply. Regulation of a *high-power switching regulator* is accomplished by push-pull switching transistors operating under control of a feedback network consisting of a pulse-width modulator and a voltage comparison amplifier, Fig. 19-14. The feedback elements control the onperiods of the switching transistors to adjust the duty cycle of the bipolar waveform (E) delivered to the output rectifier filter. Here the waveform is rectified and averaged to provide a dc output level that is proportional to the duty cycle of the waveform, varying the ontimes of the switches.

The waveforms of Fig. 19-14 provide a more detailed picture of circuit operation. The voltage comparison amplifier continuously compares a fraction of the output voltage with a stable reference voltage, V_{ref}, to produce the $V_{control}$ level for the turn-on comparator. This device compares the $V_{control}$ input with a triangular ramp waveform A, occurring at a fixed 40 kHz rate. When the ramp voltage is more positive than the control level, a turn-on signal (B) is generated. Notice that an increase or decrease in the $V_{control}$ voltage varies the width of the output pulses at B and thus the on time of the switches.

Steering logic within the modulator chip causes switching transistors Q_1 and Q_2 to turn on alternately so

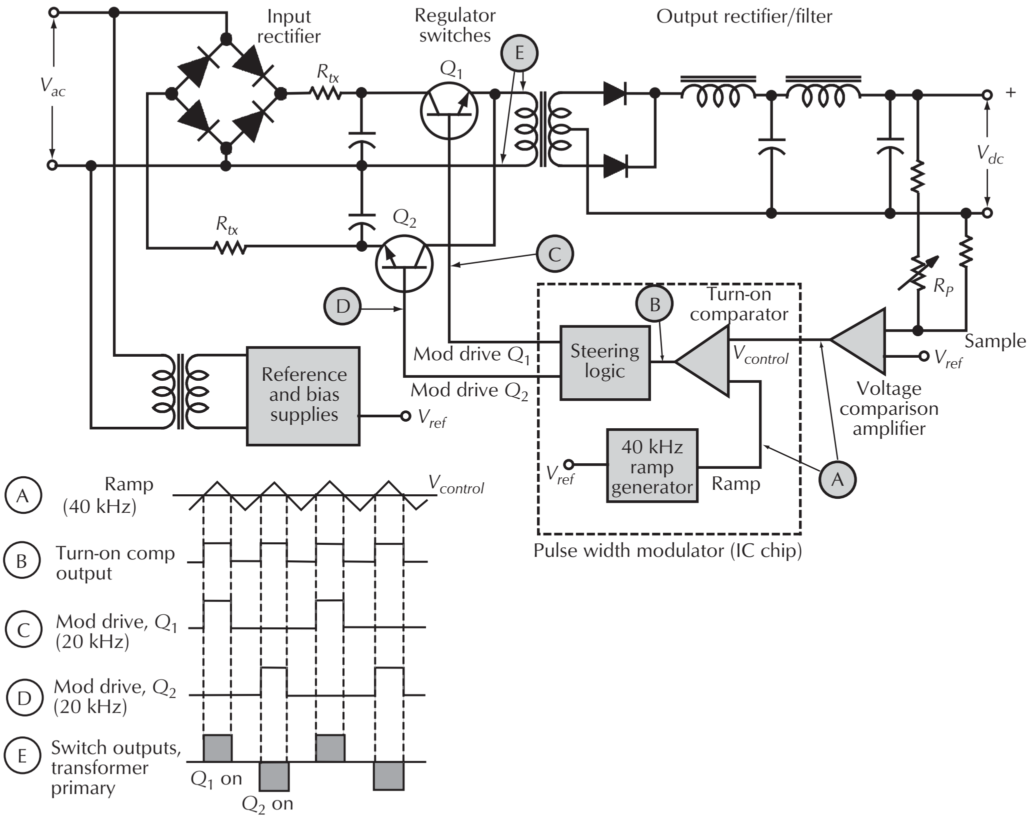

Input rectifier
Regulator switches
Output rectifier/filter
E
Q_1
R_{tx}
V_{ac}
+
V_{dc}
R_{tx}
Q_2
C
B
Turn-on comparator
D
R_P
Sample
Mod drive Q_1
Steering logic
$V_{control}$
V_{ref}
Mod drive Q_2
Reference and bias supplies
V_{ref}
Voltage comparison amplifier
V_{ref}
40 kHz ramp generator
Ramp
A
A
Ramp (40 kHz)
$V_{control}$
Pulse width modulator (IC chip)
B
Turn-on comp output
C
Mod drive, Q_1 (20 kHz)
D
Mod drive, Q_2 (20 kHz)
E
Switch outputs, transformer primary
Q_1 on
Q_2 on

Figure 19-14. Switching supply with push-pull transistors and feedback for regulation.

that each switch operates at one-half the ramp frequency or 20 kHz.

The addition of a triac preregulator and associated control circuit improves regulation and ripple. The triac is a bidirectional device and is usually connected in series with one side of the input primary. Whenever a gating pulse is received, the triac conducts current in a direction that is dependent on the polarity of the voltage across it. The goal is to control the triac so that the bridge rectifier output (dc input to the switches) is held relatively constant. This is accomplished by a control circuit that issues a phase-adjusted firing pulse to the triac once during each half-cycle of the input ac. The control circuit compares a ramp function to a rectified ac sine wave to compute the proper firing time for the triac.

Although the addition of the preregulator circuitry increases complexity, it provides three important benefits:

1. By keeping the dc input to the switches constant, it permits the use of more readily available lower voltage switching transistors.
2. The coarse preregulation it provides allows the main regulator to achieve a finer regulation.
3. Through the use of slow-start circuits, the initial conduction of the triac is controlled, providing an effective means of limiting input surge current.

19.4.4 Phase-Controlled Regulated Power Supplies

In the *phase-controlled* supply, the pass element is switched on and off at line frequency and controls the output voltage by a varying pulse width. This is most often accomplished by using an SCR as the pass element. By delaying the firing point of the SCR_1 in each cycle, the output voltage can be varied, Fig. 19-15. SCR_1 is fired by applying a voltage to the gate. The voltage is obtained by C_1 charging through R_2 and the ballast lamp. When the gate firing voltage is reached across C_1, SCR_1 fires. Once the SCR_1 is on, it remains on until its anode voltage goes to zero, which is during the second half of the cycle. When SCR_1 is on, C_1 discharges and remains discharged until the phase of the line voltage returns to zero. The rate that the C_1 charges is controlled by Q_1. When Q_1 is turned on, much of the C_1 charging current is shunted around C_1, requiring a longer time to charge C_1, thus delaying the firing of SCR_1. As the line voltage increases, the resistance of VDR_1 and VDR_2 decreases, turning Q_1 on more and thus slowing the charging rate of C_1. Since the output is a series of pulses with a high rise time of the leading edge, a filter is required on the output to smooth the dc.

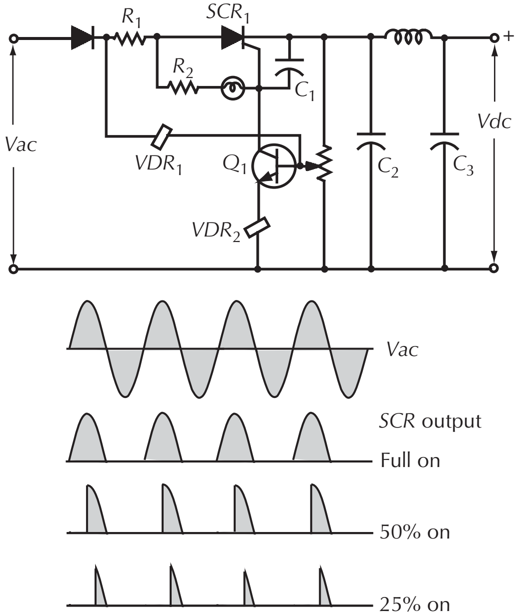


Figure 19-15. Phase controlled regulated supply.

19.5 Single IC, Power Factor Corrected, Off-Line Supply

Many off-line power supplies now include power factor correction (PFC) which reduces input current and meets regulatory requirements.[1] Normal switching power supplies that incorporate a bridge rectifier followed by bulk capacitance create harmonic currents, increasing the supply's rms input current, while contributing nothing to real power. To solve this problem a PFC preregulator and a separate controller was added to an existing design.

The Linear Technology Corporation LT®1508 (voltage mode) and LT1509 (current mode) power supply eliminate combine the PFC and a pulse width modulator (PWM) function in a single 20-pin IC.

PFC is achieved by programming the input current of a boost regulator to follow the input line voltage, resulting in a near-unity power factor compared to 0.5–to 0.7 for a typical capacitive input switcher. The architecture maintains 0.99 power factor over a 20:1 load range. Start-up is controlled by separate PFC and PWM soft start pins. The PWM Soft Start pin is held low, disabling the PWM output until the PFC stage is in regulation. The PWM will remain enabled as long as the PFC output voltage stays above 73% of its preset value (typically 280 V out of 383 V for universal input). A separate overvoltage protection pin can be connected to the output through an independent resistor divider. This ensures overvoltage protection during safety agency abnormal testing conditions, such as opening the main feedback path. The two stages are synchronized and the PWM turn-on is delayed for 50% of the oscillator cycle. This minimizes noise and conducted emission problems. 2 A peak current gate drivers and a 1.2 V optoisolator offset on the VC pin further simplify the design.

A universal input, 24 Vdc, 300 W converter using the LT1508 is shown in Fig. 19-16. Following the PFC boost preregulator is a 2-transistor forward converter that features low voltage (500 Vdc) switches, low peak currents and automatic nondissipative core reset. Under worst case conditions (low line, full power), the PFC and PWM stages have efficiencies of 90% and 92% respectively. The LT1508's low start-up current of 250 μA minimizes start-up resistor power dissipation. An overwinding on T1 provides the bootstrapped chip supply. The intermediate bus voltage of 382 V is well controlled, simplifying the post regulator and increasing capacitor holdup time compared to a typical off-line converter.

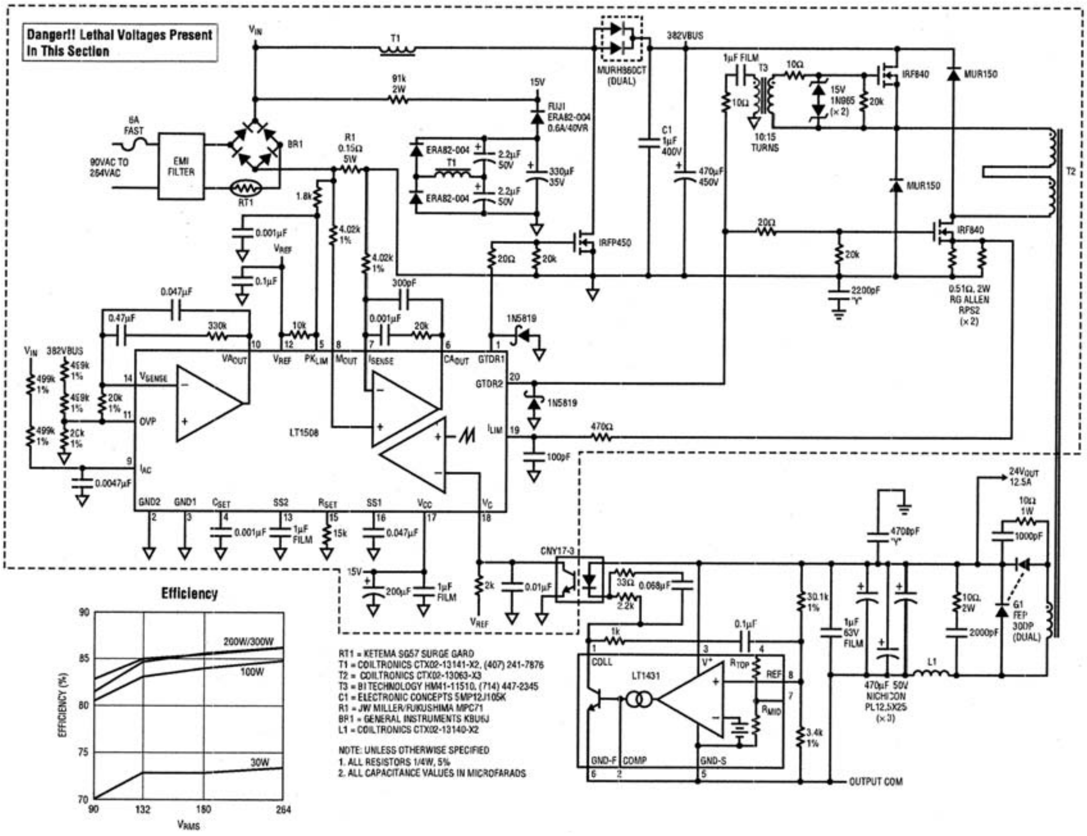
19.6 Synchronous Rectification Low-Voltage Power Supplies

Synchronous rectifiers can improve switching-power-supply efficiency, particularly in low-voltage low-power applications compared to Schottky-diode types of supplies.[2]

A synchronous rectifier is an electronic switch that improves power-conversion efficiency by placing a low resistance conduction path across the diode rectifier in a switch-mode regulator. MOSFETs or bipolar transistors and other semiconductor switches can be used.

The forward-voltage drop across a switch-mode rectifier is in series with the output voltage, so losses in the rectifier determine efficiency.

Even at 3.3 V, rectifier loss is significant. For step-down regulators with a 3.3 V output and a 12 V input, the 0.4 V forward voltage of a Schottky diode represents a typical efficiency penalty of about 12%. The losses are less at lower input voltages because the recti-


Danger!! Lethal Voltages Present In This Section
VIN
T1
382VBUS
MURH860CT (DUAL)
91k 2W
15V
FUJI ERA82-004 0.6A/40VR
6A FAST
90VAC TO 264VAC
EMI FILTER
BR1
RT1
R1 0.15Ω 5W
ERA82-004
2.2µF 50V
330µF 35V
C1 1µF 400V
470µF 450V
1µF FILM
T3
10Ω
15V 1N965 (×2)
20k
IRF840
MUR150
10:15 TURNS
T2
IRFP450
20Ω
20k
2200pF "Y"
0.51Ω, 2W RG ALLEN RPS2 (×2)
1.8k
0.001µF
VREF
0.1µF
4.02k 1%
300pF
0.047µF
0.47µF
330k
10k
1N5819
VSENSE
OVP
IAC
VAOUT
PKLIM
MOUT
ISENSE
CAOUT
GTDR1
GTDR2
ILIM
LT1508
GND2
GND1
CSET
SS2
RSET
SS1
VCC
VC
470Ω
100pF
499k 1%
20k 1%
2k 1%
0.0047µF
0.001µF
1µF FILM
15k
200µF
2k
0.01µF
CNY17-3
33Ω
0.068µF
2.2k
1k
24VOUT 12.5A
10Ω 1W
1000pF
4700pF "Y"
0.1µF
LT1431
COLL
V+
RTOP
REF
RMID
GND-F
COMP
GND-S
30.1k 1%
3.4k 1%
1µF 63V FILM
470µF 50V NICHICON PL12.5X25 (×3)
L1
10Ω 2W
2000pF
G1 FEP 30DP (DUAL)
OUTPUT COM
Efficiency
EFFICIENCY (%)
VRMS
90
85
80
75
70
132
180
264
200W/300W
100W
30W

RT1 = KETEMA SG57 SURGE GARD
T1 = COILTRONICS CTX02-13141-X2, (407) 241-7876
T2 = COILTRONICS CTX02-13063-X3
T3 = BI TECHNOLOGY HM41-11510, (714) 447-2345
C1 = ELECTRONIC CONCEPTS 5MP12J105K
R1 = JW MILLER/FUKUSHIMA MPC71
BR1 = GENERAL INSTRUMENTS KBU6J
L1 = COILTRONICS CTX02-13140-X2

NOTE: UNLESS OTHERWISE SPECIFIED
1. ALL RESISTORS 1/4W, 5%
2. ALL CAPACITANCE VALUES IN MICROFARADS

Figure 19-16. 24 V, 300 W off-line PFC supply. Courtesy Linear Technology Corporation.

fier has a lower duty cycle and thus a shorter conduction time. However, the Schottky rectifier's forward drop is usually the dominant loss mechanism.

For an input voltage of 7.2 V and an output of 3.3 V, a synchronous rectifier improves on the Schottky diode rectifier's efficiency by around 4%. As output voltage decreases, the synchronous rectifier provides even larger gains in efficiency, Fig. 19-17.

19.6.1 Diode versus Synchronous Rectifiers

In the absence of a parallel synchronous rectifier, the drop across the rectifier diode in a switching regulator, Fig. 19-18A causes an efficiency loss that worsens as the output voltage falls. The Schottky diode simple buck converter clamps the switching node, the inductor's swinging terminal, as the inductor discharges.

In the synchronous-rectifier version of Fig. 19-18B, a large N-channel MOSFET switch replaces the diode

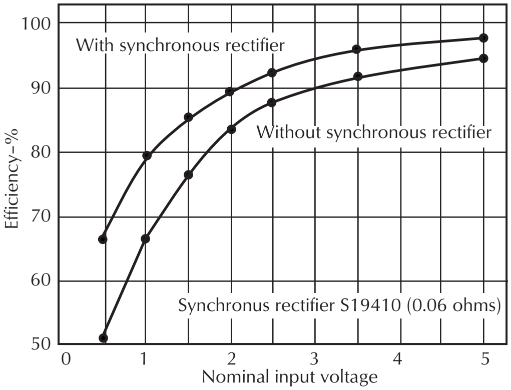


Figure 19-17. Data based on a high-performance buck switch-mode regulator and powered from a standard 7.2 V notebook-computer battery shows that the synchronous rectifier has little effect on efficiency at 5 V, but offers significant improvements at 3.3 V and below. Courtesy Maxim Integrated Products.

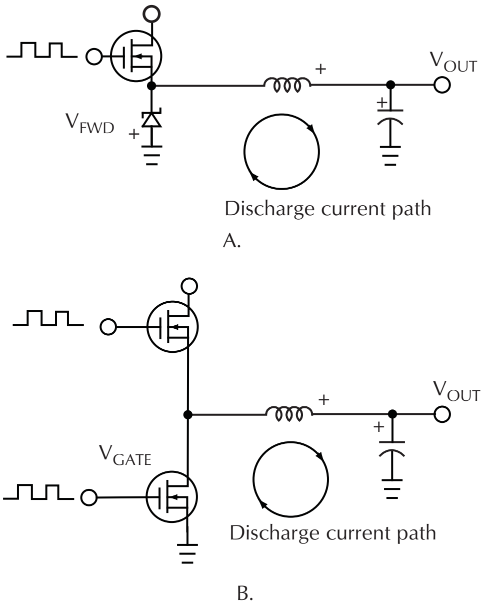


Figure 19-18. A synchronous rectifier replaces the Schottky diode in A with a low $R_{DS(ON)}$MOSFET in B. The lower-resistance conduction path improves efficiency for the 5–3.3 V 3 A converter by 3–4%. Courtesy Maxim Integrated Products.

and forms a half-bridge configuration that clamps the switching node to –0.1 V or less. The diode in Fig. 19-18A clamps the node to –0.35 V. Intuitively, losses in either type of rectifier increase with reduced output voltage. At $V_{IN}2V_{OUT}$, the rectifier voltage drop is in series with the load voltage for about half the switching period. As the output voltage falls, power lost in the rectifier becomes a greater fraction of the load power.

The basic trade-off between using diode or MOSFET rectifiers is whether the power needed to drive the MOSFET gate cancels the efficiency gained from a reduced forward-voltage drop. The synchronous rectifier's efficiency gain depends strongly on load current, battery voltage, output voltage, switching frequency, and other application parameters. Higher battery voltage and lighter load current enhance the value of a synchronous rectifier. The duty factor, which equals $1 – D$, where D equals $t_{on}/(t_{on} + t_{off})$, for the main switch, increases with the battery voltage. Also, the forward drop decreases with the load current.

The gate-drive signal is a key factor in calculating a synchronous rectifier's efficiency gain. For example, the gate loss can be reduced by using a gate drive of 5 V (as for logic-level MOSFETs) instead of the input (battery) voltage. Simply supply the gate drive from a 5 V linear regulator powered from the battery. Another method is to bootstrap the gate driver's power-supply rails from the regulator's output voltage. (This approach adds complexity in the form of a bypass switch for the initial power-up.) One must weigh the lower loss associated with reduced gate voltage against the higher $R_{DS(ON)}$ resulting from a less-enhanced MOSFET.

When comparing diode and synchronous rectifiers, note that the synchronous rectifier MOSFET doesn't always replace the usual Schottky diode. To prevent switching overlap of the high-side and low-side MOSFETs that might cause destructive cross-conduction currents, most switching regulators include a dead-time delay. The synchronous rectifier MOSFET contains an integral, parasitic body diode that can act as a clamp and catches the negative inductor voltage swing during this dead time. This diode is lossy, is slow to turn off, and can cause a 1–2% efficiency drop.

To squeeze the last percent of efficiency out of a power supply, a Schottky diode can be placed in parallel with the synchronous rectifier MOSFET. This diode conducts only during the dead time. A Schottky diode in parallel with the silicon body diode turns on at a lower voltage, ensuring that the body diode never conducts. Generally, a Schottky diode used in this way can be smaller and cheaper than the type the simple buck circuit requires, because the average diode current is low. (Schottky diodes usually have peak current ratings much greater than their dc current ratings.) It's important to note that conduction losses during the dead time can become significant at high switching frequencies. For example, in a 300 kHz converter with a 100 ns dead time, the extra power dissipated is equal to

$$I_{LOAD} \times V_{FWD} \times td \times f = 6\text{ mW} \qquad (19\text{-}22)$$

where,

f is the switching frequency

td is the dead time) for a 2.5 V, 1 W supply, which represents an efficiency loss of about 0.5%.

Light-load efficiency is a key parameter when the load spends a long time in a nearly dormant suspend mode. For the buck-type switch-mode regulators, the synchronous rectifier's control circuit has a strong influence on light-load efficiency and noise performance. The key issue for light-load or no-load conditions is the timing of the MOSFET's turn-off signal.

When load current is light, the inductor current discharges to zero, becoming discontinuous or reversing direction. There are at least three options in dealing with this problem:

1. Continue to hold the synchronous switch on until the beginning of the next cycle, allowing the inductor to reverse.
2. Completely disable the synchronous rectifier at light loads.
3. Sense the inductor current's zero crossing and shut off the synchronous rectifier on a cycle-by-cycle basis.

Each approach involves a trade-off in different areas. In the past, the option that designers widely used was holding the inductor switch on until the beginning of the next cycle which requires driving the MOSFET gates with complementary waveforms. This approach produces lower noise and allows a simple control scheme. The gate-drive signal is simply an inverted, opposite phase version of the drive signal for the high-side switch. Noise is lower because the absence of pulse skipping ensures a constant switching frequency, regardless of load. A constant, fundamental switching frequency ensures that output ripple and EMI at the harmonic frequencies won't cause havoc in the IF bands of an audio or radio system. This approach also eliminates the dead time during which a resonant-tank circuit comprising the inductor and stray capacitance at the switching node can introduce ringing.

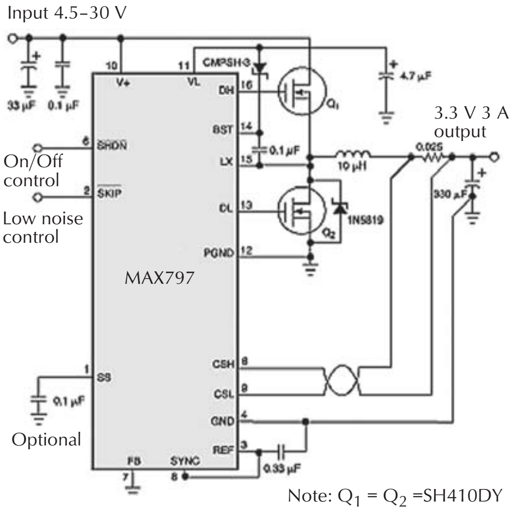
Unfortunately when the inductor current reverses, the synchronous rectifier pulls current from the output. The circuit replaces this lost output energy during the next half cycle. However, at the beginning of the cycle when the high-side switch turns on, the circuit transfers the inductor energy stored during the earlier current reversal to the input-bypass capacitor.

This action resembles perpetual motion, in which energy shuttles between the input and output capacitors. As energy shuttles back and forth, the circuit dissipates power in all its parasitic resistances and switching inefficiencies, so additional energy is necessary to maintain the shuttling action. The most obvious consequence is a high no-load supply current of typically 5 mA for the 2.5 V, 1 W circuit.

The second option, turning off the synchronous rectifier entirely at light loads, offers simplicity and low quiescent supply current. This method can be implemented in conjunction with a pulse-skipping operation, governed by a light-load pulse-frequency-modulation (PFM) control scheme. Whenever the circuit goes into its light-load pulse skipping mode, the circuit disables the synchronous rectifier that lets an accompanying parallel Schottky diode do all the work. Disabling the synchronous rectifier prevents the reversal of inductor current, and the problem of shuttling energy back and forth does not arise.

The final option, sensing the inductor current's zero crossing and quickly latching the synchronous rectifier off, turns off the synchronous rectifier on a cycle-by-cycle basis. This method provides the highest light-load efficiency, because the synchronous rectifier does its job without allowing the inductor current to reverse. But, to be effective, the switching-regulator IC's current-sense amplifier that monitors the inductor current must combine high speed with low power consumption.

A logic-control input can shift the synchronous-rectifier operation from the complementary-drive option to the off-at-zero option, Fig. 19-19. When low, "SKIP" allows normal operation: The circuit employs pulse-width modulation (PWM) for heavy loads and automatically switches to a low-quiescent-current pulse-skipping mode for light loads. When high, "SKIP" forces the IC to a low-noise fixed-frequency PWM mode, regardless of the load. Also, applying a high level to "SKIP" disables the IC's zero-crossing detector, allowing the inductor current to reverse direction, which suppresses the parasitic resonant LC tank circuit.



Note: $Q_1 = Q_2$ =SH410DY

Figure 19-19. A N-channel buck regulator which has a low-noise logic-control input that adjusts the synchronous rectifier's timing on the fly. Courtesy Maxim Integrated Products.

Another issue related to a synchronous rectifier's gate-drive timing is the cross regulation of multiple outputs obtained using flyback windings. Placing an extra winding or a coupled inductor on a buck regulator's inductor core can provide an auxiliary output voltage for the cost of a diode, a capacitor, and a little wire, Fig. 19-20.

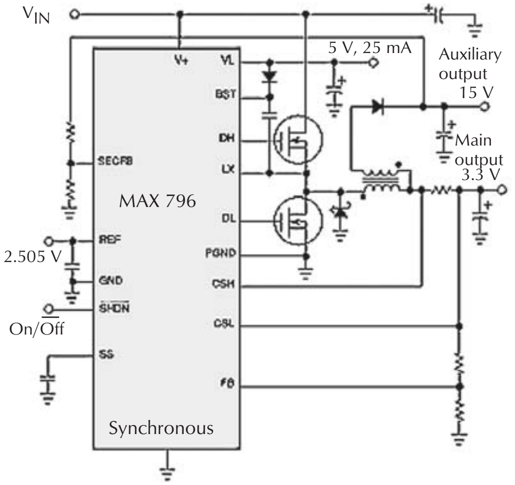


Figure 19-20. A feedback input for the secondary winding (SECFB) greatly improves the cross regulation for multiple outputs under conditions of light primary loading or low I/O differential voltage. Courtesy Maxim Integrated Products.

Normally, the coupled-inductor flyback trick in Fig. 19-20 stores energy in the core when the high-side switch is on and discharges some of it through the secondary winding to an auxiliary 15 V output when the synchronous rectifier's low-side switch is on. During discharge, the voltage across the primary is equal to $V_{OUT} + V_{SAT}$, where V_{OUT} is the main output and V_{SAT} is the synchronous rectifier's saturation voltage. Therefore, the secondary output voltage equals the primary output times the turns ratio.

Unfortunately, if the synchronous rectifier turns off at zero current and the primary load is light or nonexistent, the 15 V output sags to ground because the core stores no energy at this time. If the synchronous rectifier remains on, the primary current can reverse and let the transformer operate in the forward mode, providing a theoretically infinite output-current capability that prevents the 15 V output from sagging. Unfortunately, quiescent supply current suffers a great deal.

However, the circuit in Fig. 19-20 achieves excellent cross regulation with no penalty in quiescent supply current. A second, extra feedback loop senses the 15 V output. If this output is in regulation, the synchronous rectifier turns off at zero current as usual. If the output drops below 13 V, the synchronous rectifier remains on for an extra microsecond after the primary current reaches zero, so the 15 V output can deliver hundreds of milliamps even with no load on the main 5 V output. This scheme also provides a better 15 V load capability at low values of $V_{IN} - V_{OUT}$, which becomes important if the input voltage drops.

19.6.2 Secondary-Side Synchronous Rectifiers

Multiple synchronous rectifiers on the secondary windings can replace the usual high-voltage rectifier diodes in multiple-output nonisolated applications, Fig. 19-21. This substitution can dramatically improve load regulation on the auxiliary outputs and often eliminates the need for linear regulators, which are otherwise added to increase the output accuracy. The MOSFET must be selected with a breakdown rating high enough to withstand the flyback voltage, which can be much higher than the input voltage. Tying the gates of the secondary-side MOSFETs directly to the gate of the main synchronous MOSFET (the DL terminal) provides the necessary gate drive.

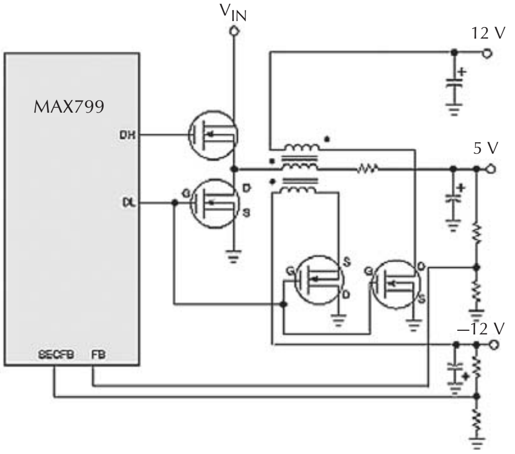


Figure 19-21. Coupled-inductor secondary outputs can benefit from synchronous rectification. To accommodate negative auxiliary outputs, swap the secondary-side MOSFET's drain and source terminals. (For clarity, this simplified schematic omits most of the ancillary components needed to make the switching regulator work.) Courtesy Maxim Integrated Products.

Another neat trick enables a synchronous rectifier to provide gate drive for the high-side switching MOSFET. Tapping the external switching node to generate a gate-drive signal higher than the supply voltage enables the use of N-channel MOSFETs for both switches in a synchronous-rectifier buck converter. Compared to P-channel types, N-channel MOSFETs have many advantages, because their superior carrier mobility confers a near 2:1 improvement in gate capacitance and on-resistance.

A flying-capacitor boost circuit provides the high-side gate drive, Fig. 19-22. The flying capacitor is in parallel with the high-side MOSFET's gate-source terminals. The circuit alternatively charges this capacitor from an external 5 V supply through the diode and places the capacitor in parallel with the high-side MOSFET's gate-source terminals. The charged capacitor then acts as supply voltage for the internal gate-drive inverter, which is comparable to several 74HC04 sections in parallel. Biased by the switching node, the inverter's negative rail rides on the power-switching waveform at the LX terminal.

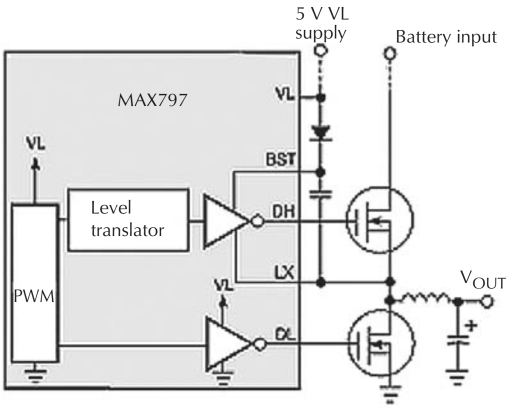


Figure 19-22. Driven by the switching node (the left end of the inductor), the capacitor between BST and LX provides an elevated supply rail for the upper gate-drive inverter. Courtesy Maxim Integrated Products.

A flying capacitor then acts as supply voltage for the internal gate-drive inverter, which is comparable to several 74HC04 sections in parallel. Biased by the switching node, the inverter's negative rail rides on the power-switching waveform at the LX terminal.

The synchronous rectifier is indispensable to the Fig. 19-22 gate-drive boost supply. Without this low-side switch, the circuit may not start at initial power-up. When power is first applied, the low-side MOSFET forces the switching node to 0 V and charges the boost capacitor to 5 V.

Synchronous rectifiers can be incorporated in the boost and inverting topologies. The boost regulator in Fig. 19-23 employs an internal pnp synchronous rectifier in the active rectifier block. Boost topologies require the rectifier in series with V_{OUT}, so the IC connects the pnp collector to the output and the emitter to the switching node. The rectifier control block's fast comparator detects whether the rectifier is forward- or reverse-biased and drives the pnp transistor on or off accordingly. When the transistor is on, an adaptive base-current control circuit keeps the transistor on the edge of saturation. This condition minimizes the efficiency loss due to base current and maintains high switching speed by minimizing the delay due to stored base charge.

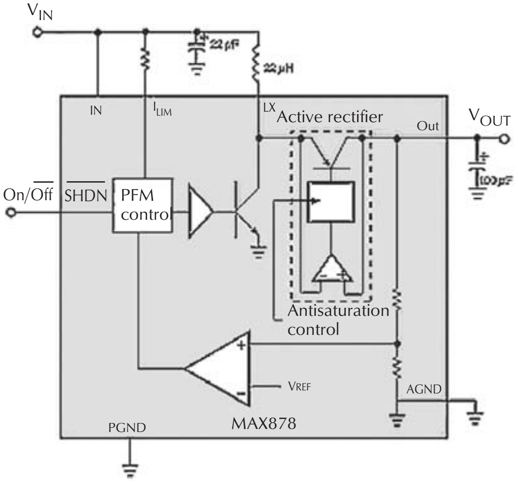


Figure 19-23. The internal synchronous rectifier in this boost regulator, the active rectifier, replaces the Schottky rectifier often used at that location. Courtesy Maxim Integrated Products.

An interesting side benefit of the pnp synchronous rectifier is its ability to provide both step-up and step-down action. For ordinary boost regulators, the input voltage range is limited by an input-to-output path through the inductor and the diode. (This unwanted path is inherent in the simple boost topology.) Thus, if V_{IN} exceeds V_{OUT}, the conduction path through the rectifier can drag the output upward, possibly damaging the load with overvoltage.

The pnp-rectifier circuit in Fig. 19-23 operates in switch mode, even when V_{IN} exceeds V_{OUT}, with the active rectifier acting as the switch. This action is more akin to a regulating charge pump than to a buck regulator, because the buck mode of operation requires a second switch on the high side.

Inverting-topology regulators that generate negative voltages, sometimes called buck-boost regulators, are useful applications for synchronous rectification. Like the boost topology, the inverting topology connects the synchronous rectifier in series with the output rather than to ground, Fig. 19-24. In this example, the synchronous switch is an N-channel MOSFET with its source tied to the negative output and its drain tied to the switching node.

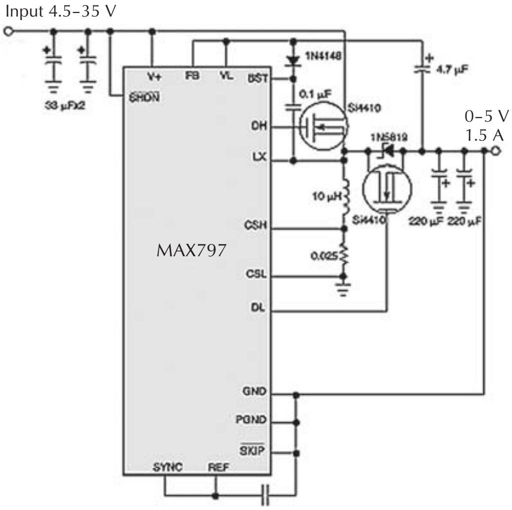


Figure 19-24. The inverting topology requires that the synchronous switch be in series with the output. Courtesy Maxim Integrated Products.

The circuit tricks the resulting 300 kHz buck regulator into performing as an inverting-topology switcher by connecting the IC's GND pin to the negative output voltage instead of circuit ground. This switching regulator's efficiency of about 88% exceeds that of comparable asynchronous-rectifier supplies by 4%.

19.7 Converters

A *converter* changes low-voltage dc to high-voltage dc. Basically, a dc-to-dc converter consists of a dc source of potential (generally a battery) applied to a pair of switching transistors. The transistors convert the applied dc voltage to a high-frequency ac voltage. The ac voltage is then transformed to a high voltage that is rectified to dc again and filtered in the conventional manner. Power supplies of this nature are often used for a source of high voltage, where the usual ac line voltage is not available.

19.8 Inverters

An *inverter* converts direct current to alternating current. Inverters are used in applications where the primary source of power is direct current. Because direct current cannot be transformed, it is converted to alternating current so that alternating current output from the inverter may be applied to a transformer to supply the desired voltage.

An inverter operates much like the switching circuit and transformer section of a converter. In Fig. 19-25, R_1 and R_2 assure that the oscillator (switch) will start. T_1 is a saturable base-drive transformer that determines the drive current to turn on Q_1 or Q_2. T_2 is a non-saturable transformer; therefore, collector current through Q_1 and Q_2 is dependent upon load. Base resistors R_b are current limiting resistors. By adding a rectifier and filter section, this inverter can be changed to a converter.

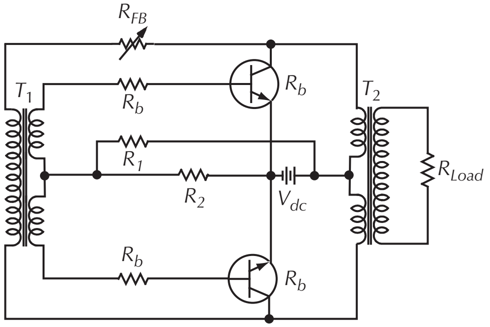


Figure 19-25. Two-transistor, two-transformer, push-pull inverter that uses a resistive voltage-divider network to provide starting bias.

19.9 Ultra Capacitor (UPS)

A backup power supply for medical computers manufactured by Ram Technologies utilizes ultra capacitor technology. The model 8000 Ultra UPS module contains the charge and discharge circuitry to ensure high-efficiency energy transfer, Fig. 19-26. The proprietary patent pending module is designed to directly interface with RAM Technologies line of ATX/SFX medical-grade power supplies. The unit can be modified by Ram Technologies to operate with other sensitive and/or life-threatening devices. The module may be expanded by adding additional ultra capacitor modules. The base module contains 8000 J of energy; expansion modules also contain 8000 J of energy. Any number of additional modules can be added to increase load capabilities. Fig. 19-27 shows a typical installation in a computer.

The module's input voltage is +12 Vdc and has an efficiency of >90%. Charge time is 2 minutes for each 8 kJ module.

The maximum output current is 30 A at 12 Vdc. Run time is

$$Run\ time = \frac{number\ of\ modules \times 133}{dc\ load} \tag{19-23}$$

where,
Run time is in minutes,
dc load is in watts.

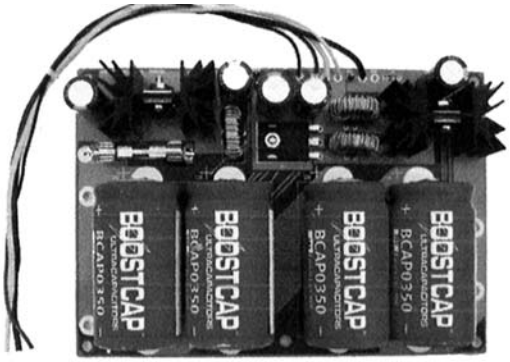

BOOSTCAP
BCAP0350
BOOSTCAP
BCAP0350
BOOSTCAP
BCAP0350
BOOSTCAP
BCAP0350

Figure 19-26. Ultra capacitor UPS power supply module. Courtesy Ram Technologies LLC.

19.9.1 Ultra Capacitor Characteristics

Ultra Capacitors do not degrade in time as batteries do so reliability is high.

UCs have a capacity one million times that of a standard electrolytic capacitors. This is accomplished by depositing carbon on an aluminum substrate which can be etched it give dramatically increased surface area and low impedance. A typical UC 30 mm × 50 mm can have up to 400 F of capacitance with a working voltage of 2.5 V.

The energy stored in a capacitor is $CV^2/2$. To have the same energy density as current Lithium Ion batteries, the working voltage for a given size would have to be increased to 5 V. This would yield 4 times the energy density of current UC technology and make UCs the ultimate energy storage device. It is just a matter of time before technology reveals materials capable of having dielectrics in the 5 V or higher range and making batteries as we know them today obsolete.

Charging. Ultra Capacitors have extremely low internal impedance so power limited charging must be used to avoid overloading the charger. UCs are also sensitive to voltage. They must be operated below their rating to avoid destruction. Designers tend to charge them as close as possible to ther maximum voltage to extract the maximum energy from them.

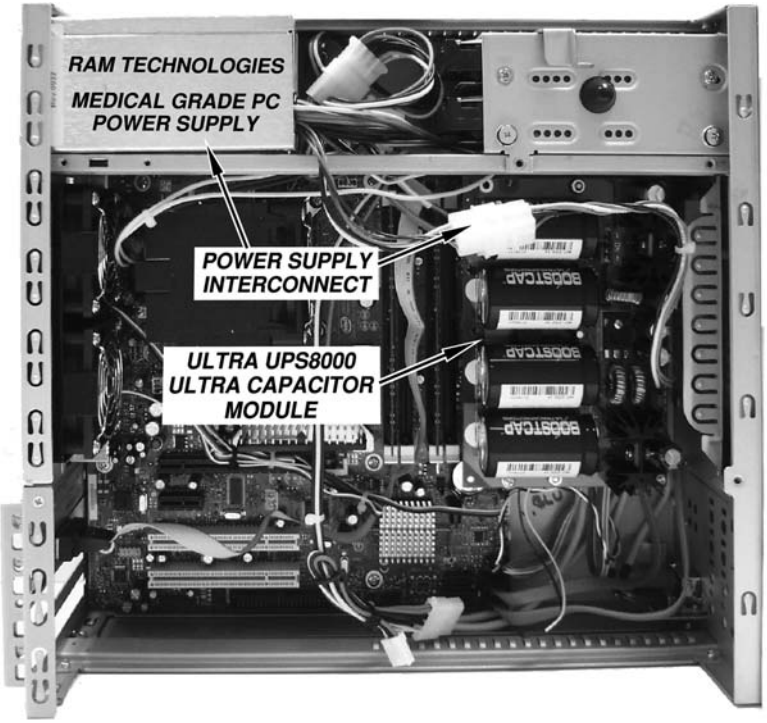

RAM TECHNOLOGIES
MEDICAL GRADE PC
POWER SUPPLY
POWER SUPPLY
INTERCONNECT
ULTRA UPS8000
ULTRA CAPACITOR
MODULE
BOOSTCAP
BOOSTCAP
BOOSTCAP

Figure 19-27. Typical installation of a ultra capacitor UPS module in a computer. Courtesy Ram Technologies LLC.

Discharging: Unlike batteries Ultra Capacitors store their energy over their entire voltage range which makes the design of the boost complex due to the wide input voltage range and overall efficiency.

19.10 Batteries

Batteries offer a means of producing a smooth, ripple-free, hum-free, portable power supply. A battery's capacity is rated in ampere-hours (Ah). Three facts about batteries are:

1. An ampere-hour can be a 1 A drain for 1 h, 0.5 A drain for 2 h, or 2 A drain for 0.5 h.
2. A 12 V liquid battery is generally considered completely discharged when its voltage reaches 10.5 V.
3. Batteries for cycling service—i.e. powering amplifiers etc.— are normally rated with:
    - A 20 h discharge rate, which means a 20 Ah battery will deliver 1 A for 20 h and a 100 Ah battery will deliver 5 A for 20 h.
    - A reserve capacity stated in minutes for a 25 A discharge rate.
    - Discharging batteries below 50% shortens their life.

A cell or battery is an electrochemical system that converts chemical energy into electrical energy. When the chemical action is reversible, the battery is a secondary or rechargeable system.

To be rechargeable, the positive and negative electrodes of a battery must be capable of being converted back to their original state following a discharge. Thus, the battery must be electrically recharged by reversing the process that occurred during its discharge cycle.

19.10.1 Temperature Effects

The standard rating for batteries is at 25°C (77°F). Battery capacity is reduced at lower temperature. At freezing, Ah capacity is reduced to 80%. At –27°C (–22°F), Ah capacity drops to 50%. At 122°F, capacity is increased by 12%.

Battery charging voltage is also affected by temperature. It will vary from about 2.74 V/cell (16.4 V) at –40°C (–40°F) to 2.3 volts per cell (13.8 V) at 50°C (122°F).

Temperature also affects battery life. While battery capacity is reduced by 50% at –22°F, battery life increases about 60%. Battery life is reduced at higher temperatures. In fact, for every 8.3°C (15°F) over 25°C (77°F), battery life is cut in half. This holds true for all types of lead-acid batteries, sealed, gelled, and AGM.

19.10.2 Cycles versus Battery Life

A battery cycle is one complete discharge and recharge cycle and is often considered a discharge from 100% to 20%, and then recharged back to 100%. Other ratings for depth of discharge (DOD) cycles are 10%, 20%, and 50%.

Battery life is directly related to how deep the battery is cycled each time. If a battery DOD is 50% every cycle, it will last twice as long as if the DOD is 80%. If the DOD cycle is only 10%, it will last about five times as long as one cycled to 50%. A 50% DOD is usually recommended. A battery that has a DOD cycle of 5% or less usually does not last as long as one cycled down 10% because at very shallow cycles, the lead dioxide tends to build up in clumps on the positive plates rather in an even film.

19.10.3 Battery Voltage

All lead-acid batteries supply about 2.14 volts per cell (V/cell), or 12.6–12.8 V for a 12 volt battery when fully charged. Batteries that are stored for long periods will eventually self-discharge. This varies with battery type, age, and temperature. Self-discharge can range from 1–15% per month. Batteries should never be stored in a partly discharged state for a long period of time. A float charge should be maintained if they are not used.

19.10.4 State of Charge

State of charge, or conversely, the depth of discharge (DOD), can be determined by measuring the voltage and/or the specific gravity of the acid with a hydrometer. Voltage on a fully charged battery is 2.12–2.15 V/cell, or 12.7 V for a 12 volt battery. At 50% DOD the voltage is 2.03 V/cell, and at 0% DOD it is 1.75 V/cell or less. Specific gravity is 1.265 for a fully charged cell and 1.13 or less for a totally discharged cell. Many batteries are sealed, therefore, hydrometer reading cannot be taken.

19.10.5 False Capacity

A battery can meet all the tests for being at full charge, yet be lower than its original capacity because the plates are damaged, sulfated, or partially gone from long use.

In this case it acts like a battery of much smaller size.

19.10.6 Ampere-Hour Capacity

Deep cycle batteries are rated in ampere hours (Ah). An Ah is a 1 A drain for 1 h, 10 A for 0.1 h, etc. It is calculated with the equation A × h. Drawing 20 A for 20 min would be 20 A × 0.333 h, or 6.67 Ah. The accepted Ah rating time period for batteries used in solar electric and backup power systems and for nearly all deep cycle batteries is the 20 hour rate. This is defined as the battery being discharged to 10.5 V over a 20 h period while the total actual Ah it supplies is measured.

Amp-hours are specified at a particular rate because of the Peukert effect. The Peukert value is directly related to the internal resistance of the battery. The higher the internal resistance, the higher the losses while charging and discharging, especially at higher currents. The faster a battery is discharged, the lower the Ah capacity. Conversely, if it is drained more slowly, the Ah capacity is higher.

19.10.7 Battery Charging

Batteries can be charged by constant current or constant voltage. When charged by the constant-current method, care must be taken to eliminate the possibility of overcharging; therefore, the condition of the battery should be known before charging so that the charger can be removed when the ampere-hour rate of the battery is met.

Charging with the constant voltage method reduces the possibility of overcharging. With the constant voltage method, charge current is high initially and tapers off to a trickle charge when the battery is fully charged. Two requirements must be met when using the constant-voltage method:

- The charging voltage must be stable and set to 2.4 V per cell for a lead-acid battery and 2.30 V per cell for a gel cell battery. Gel cell open-circuit voltage is 2.12 V per cell.
- A current-limiting circuit must be employed to limit charge current when the battery is fully discharged.

A good battery charger charges in three steps. In the first stage, charge current is at the maximum safe rate the batteries will accept until the voltage rises to 80–90% of full charge level. Voltages at this stage typically range from 10.5–15 V. There is no correct voltage for bulk or charge charging, but there may be limits on the maximum current that the battery and/or wiring can accept.

In the second stage, accept, the voltage remains constant and current gradually tapers off as internal resistance increases during charging. Voltages are typically 14.2–15.5 V.

After batteries reach full charge, the third stage charging voltage is reduced to a lower level, 12.8–13.2 V, to reduce gassing and prolong battery life. This is often referred to as a *maintenance, float,* or *trickle charge*, since its main purpose is to keep an already charged battery from discharging, Fig. 19-28. An ideal charging state table is shown in Table 19-4.

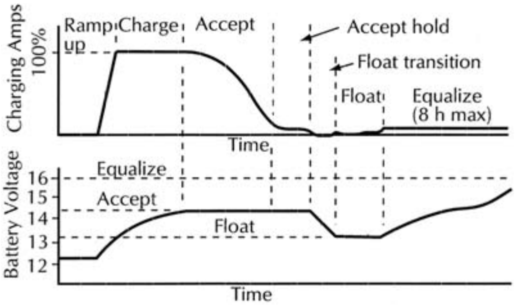


Figure 19-28. Ideal charge curve.

Table 19-3. Ideal Charging State

Cycle	Voltage	Current
Charge	12.0–14.3 Rising	Maximum
Accept	14.4 Constant	Falling
Float	13.5 Constant	Small (<2% capacity)
Equalize	13.2–16.0 Rising	Constant until 16.0 V

PWM, or *pulse width modulation* is sometimes used as a float or trickle charge. In PWM chargers, the controller circuit senses small voltage drops in the battery and delivers short charging cycles (pulses) to the battery. This may occur several hundred times per minute and is called *pulse width* because the width of the pulses varies from a few microseconds to several seconds.

Most flooded batteries should be charged at no more than the C/8 rate for any sustained period. C/8 is the battery capacity at the 20 hour rate divided by 8. For a 220 Ah battery, this would equal 26 A. Gelled cells should be charged at no more than the C/20 rate, or 5% of their amp-hour capacity. AGM batteries can be charged at up the C × 4 rate, or 400% of the capacity for the bulk charge cycle.

Lead acid batteries require 15.5 V for 100% charge. When the charging voltage reaches 2.583 V/cell, charging should be stopped or reduced to a trickle charge. Flooded batteries must bubble (gas) to insure a full charge, and to mix the electrolyte. Float voltage for flooded batteries should be 2.15–2.23 V/cell, or 12.9–13.4 V for a 12 volt battery. At higher temperatures, over 85°F, charge voltage should be reduced to 2.10 V/cell. Float and charging voltages for gelled batteries are usually about 0.2 V less than for flooded batteries.

19.10.7.1 Equalizing

The equalize cycle puts a controlled overcharge to remove lead sulfate from the plates that is not removed during the normal charging of the battery. Flooded battery life can be extended if an equalizing charge is applied every 10 to 40 days. This is a charge that is about 10% higher than the normal full charge voltage, and is applied for 2–16 h to be sure that all cells are equally charged and the gas bubbles mix the electrolyte. If the liquid in the cells is not mixed, the electrolyte becomes stratified, creating a strong solution at the top and weak solution at the bottom of the cell. AGM and gelled batteries should be equalized a maximum of two to four times a year.

19.10.7.2 Charging Voltage versus Temperature

Battery charging is sensitive to temperature. As the ambient temperature decreases, the charging voltage must be increased, Table 19-4.

Table 19-4. Temperature Compensation for Various Types of Batteries

Temp		Liquid		Gel—Std.		Gel—Fast		AGM	
°F	°C	Accept	Float	Accept	Float	Accept	Float	Accept	Float
120	49	12.5	12.5	13.0	13.0	13.0	13.0	12.9	12.9
110	43	13.6	12.7	13.5	13.0	14.0	13.4	13.9	12.9
100	38	13.8	12.9	13.7	13.2	14,1	13.5	14.0	13.0
90	32	14.0	13.1	13.8	13.3	14.2	13.6	14.1	13.1
80	27	14.2	13.3	14.0	13.5	14.3	13.7	14.2	13.2
70	21	14.4	13.5	14.1	13.6	14.4	13.8	14.3	13.3
60	16	14.6	13.7	14.3	13.8	14.5	13.9	14.4	13.4
50	10	14.8	13.9	14.4	13.9	14.6	14.0	14.5	13.5
40	5	15.0	14.1	14.6	14.1	14.7	14.1	14.6	13.6
30	-1	15.2	14.3	14.7	14.2	14.8	14.2	14.7	13.7

19.10.7.3 State of Charge

Table 19-5 shows no-load typical voltages versus state of charge for a 12 V battery. These voltages are for batteries that have been at rest for 3 hours or more. Note the large voltage drop in the last 10%.

Table 19-5. No Load Voltage versus State of Charge for a 12 V Battery

State of Charge	12 Volt battery	Volts per Cell
100%	12.7	2.1
90%	12.5	2.1
80%	12.4	2.1
70%	12.3	2.1
60%	12.2	2.0
50%	12.1	2.0
40%	11.9	2.0
30%	11.8	2.0
20%	11.6	1.9
10%	11.3	1.9
0%	10.5	1.8

19.10.7.4 Internal Resistance

All batteries have internal resistance that causes the battery voltage to fluctuate with the load. To calculate the internal resistance of a single cell or battery, the open-circuit voltage V_1 is measured using a voltmeter with an internal resistance of at least 1000 Ω/V. The battery or cell is then loaded with resistor R_1, and the voltage V_2 across the resistor is measured. R_1 should be at least 10 times the battery resistance. The current through the resistor R_1 is

$$I = \frac{V_2}{R_1}. \tag{19-24}$$

The internal resistance, R_i, of the battery may now be calculated using

$$R_i = \frac{V_1}{I} - R_1 . \tag{19-25}$$

19.10.8 Lead-Acid Batteries

The lead-acid storage battery was invented by Gaston Planté in 1860 and is one of the most widely used forms of battery power. The principal drawback to this type of battery has been the liquid electrolyte and the fumes given off when charging and discharging. Today the sealed lead-acid battery may take its place with other rechargeable batteries such as the nickel- cadmium battery. Since small amounts of gas may be generated in any battery during the charge or discharge cycle, lead-acid batteries are vented so that the gas but not the electrolyte escapes.

Lead-acid cells are normally 2.1 V and are easily connected in series to produce 6 V and 12 V automotive types, 24 V aircraft types and 36 V types for golf carts, etc. Lead-acid batteries, because of their availability, high Ah ratings, and ability to be connected in series, work well powering sound systems in the field.

The type and amount of charge determine the condition of the cell. If a lead-acid battery is overcharged, excessive water consumption and hydrogen evolution result, while constant undercharging results in a battery with less and less capacity.

The recharge factor (RF) is defined as the charge Ah divided by the previous discharge Ah. The RF must always be greater than 1 to bring the battery back to capacity. The actual RF is between 1.04 and 1.20, with the sealed lead acid batteries requiring less than the standard vented type. Fig. 19-29A shows the state of charge (SOC) achieved versus the RF for a lead-acid battery. Fig. 19-29B shows the SOC versus the RF after a number of cycles. Note that the battery rapidly loses its capacity if it is not overcharged—i.e., more is put in than is taken out.

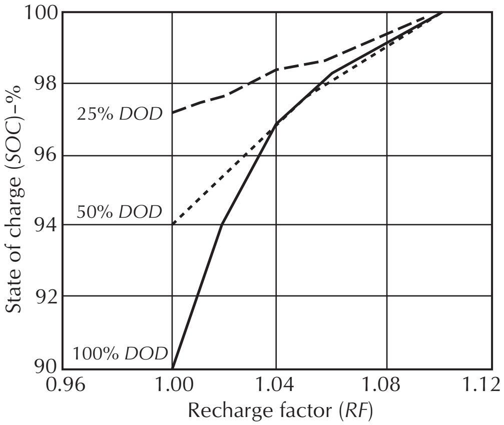
The use of a trickle charger with a storage battery shortens the life of the battery because of overcharging. Trickle chargers should only be used when it is impractical to charge a battery by other means. A practical approach to the problem is to adjust the charging voltage to a value between 2.15 V and 2.17 V per cell.

A better, but more elaborate, method is to check the specific gravity of the cells over a period of several months and adjust the charging voltage to a value where the specific gravity is maintained at 1.250. Compensation must be made for temperature changes when reading the specific gravity. Four gravity points are added to the reading for every 10°F (5°C) the electrolyte is above a temperature of 80°F (27°C).

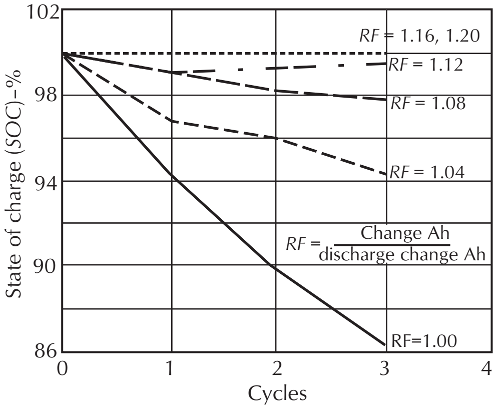
The freezing point of a battery electrolyte depends on the specific gravity of the electrolyte, Table 19-6. Lead acid batteries freeze when in a discharged state, so it is imperative that they be kept fully charged when in subfreezing temperatures. If a storage battery is left in a discharged condition for any length of time, the plates may be damaged due to sulfation.

Table 19-6. Effect of Specific Gravity on Freezing Point of a Battery

Specific Gravity	Freezing Point	Specific Gravity	Freezing Point
1.275	−85°F	1.175	+4°F
1.250	−62°F	1.150	+5°F
1.225	−35°F	1.125	+13°F
1.220	−16°F	1.100	+19°F



A. A typical state of charge for lead-acid battery versus the recharge factor.



B. State of charge versus recharge factor after three discharge cycles of a lead-acid battery.

Figure 19-29. Lead acid battery.

19.10.9 Lead-Dioxide Batteries

A *lead-dioxide battery* is a gelled electrolyte, maintenance-free type that exhibits high capacity and long life when properly applied and charged. To prevent electrolyte movement in the battery, the electrolyte in sealed batteries is immobilized by the use of a gelling agent that stores the electrolyte in highly porous separators. With this construction, the loss of water is minimized.

The terminal voltage of each cell is approximately 2.12 V. The cell voltage is higher for a battery that has just been taken off charge, but in all instances it should adjust to about 2.12 V after a period of time.

Gel/Cells comes as a type A or type B. The type A Gel/Cell is conservatively designed for 4–6 years of continuous charging in standby power applications. During this period over 100 normal discharge/charge cycles can be expected. Even more are obtained if only minor discharges are experienced. The *end of life* is actually determined by when the equipment will no longer perform its required function. Since the battery may still have 40–60% of its initial capacity, the service life may be much longer.

The type A Gel/Cell has near its full nominal capacity upon shipment from the factory. Type A cells are used for alarm systems, memory standby, etc. where they are normally in a standby mode.

The type B Gel/Cell is designed to provide 3–5 years of service in standby power applications or 300–500 normal discharge-charge cycles in portable power applications.

As the battery is discharged, the terminal voltage will slowly decrease. For instance, when the rated capacity of the battery is removed over a 20 h period, the terminal voltage would decrease to 1.75 V per cell. These batteries are rated at a 20 h current rate and room temperature. This means a 2.6 Ah battery would put out 0.13 A for 20 h. This does not mean, however, that it will put out 2.6 A for 1 h (it would put out about 1.7 A for 1 h).

Lead-dioxide batteries can be charged by the constant-current or constant-voltage method. The constant-current method is used when charger cost is the primary consideration. The battery is forced to receive a constant amount of current regardless of its needs. While charger component economy is achieved, it is sometimes done at the expense of recharge time or service life if the current is not properly set. Trickle charging current ranges from 0.5–2.0 mA per rated Ah capacity of the battery.

When charging with the constant-voltage method, a voltage of 2.25–2.30 V per cell should be used. To maintain the battery at 100°F (38°C), a voltage of 2.2 V/cell is required, while at 30°F (0°C), 2.4 V/cell is required.

19.10.10 Absorbed Glass Mat Batteries

Absorbed glass mat batteries (AGM) are sealed batteries that can be operated in any position. AGM was developed to provide increased safety, efficiency, and durability. In AGM batteries the acid is absorbed into a very fine glass mat that is not free to slosh around. The plates are kept only moist with electrolyte, so gas recombination is more efficient (99%). The AGM material has an extremely low electrical resistance so the battery delivers high power and efficiency. AGM batteries offer exceptional life cycles.

The plates in an AGM battery may be flat like wet cell lead-acid batteries, or they may be wound in a tight spiral. Their construction also allows for the lead in their plates to be purer as they no longer need to support their own weight. AGM batteries have a pressure relief valve that activates when the battery is recharged at voltage greater than 2.30 V/cell. In cylindrical AGM batteries, the plates are thin and wound into spirals so they are sometimes referred to as spiral wound.

AGM batteries have several advantages over both gelled and flooded, at about the same cost as gelled:

- All the electrolyte (acid) is contained in the glass mats so they cannot spill or leak, even if broken. Since there is no liquid to freeze and expand, they are practically immune from freezing damage.
- Most AGM batteries are recombinant—i.e., the oxygen and hydrogen recombine inside the battery. Using the gas phase transfer of oxygen to the negative plates to recombine them back into water while charging prevents the loss of water through electrolysis. The recombining is typically 99+% efficient.
- AGM batteries have a self-discharge of 1–3% per month.
- AGM batteries do not have any liquid to spill, and even under severe overcharge conditions, hydrogen emission is far below the 4% max, specified for aircraft and enclosed spaces.
- The plates in AGM's are tightly packed and rigidly mounted so they withstand shock and vibration.

19.10.10.1 A Comparison of the Three Types of Deep Cycle Batteries

Safety. Batteries can be dangerous. They store a tremendous amount of energy, create explosive gas during charge and discharge, and contain dangerous chemicals. Both gel and AGM batteries are sealed batteries that use recombinant gas technology. AGM is more efficient in the AGM process and completes its gas recombination near the plates. Gel recombinant gas batteries should incorporate automatic temperature-compensated voltage regulators to prevent explosions associated with their overcharging. Flooded batteries will spew acid, will definitely spill and leak if tipped over, and they generate dangerous and noxious explosive gases. AGM batteries are best at protecting both equipment and passengers.

Longevity. All batteries die. The number of cycles it takes to kill them is a function of the type and quality of the battery. When cycled between 25% and 50% depth of discharge (recommended deep cycle use), AGM batteries will normally outlast the other two types.

Durability. Some battery designs are simply more durable than others are. They are more forgiving in abusive conditions —i.e., they are less susceptible to vibration and shock damage, over-charging, and deeper discharge damage. Gel acid batteries are the most likely to suffer irreversible damage from overcharging. Flooded acid batteries are the most likely to suffer from internal shorting and vibration damage. AGM batteries are usually more durable and can withstand severe vibration, shocks, and fast charging.

Efficiency. Internal resistance of a battery denotes its overall charge/discharge efficiency and its ability to deliver high cranking currents without significant drops in voltage and is a measure of how well it has been designed and manufactured. Internal resistance in NiCad batteries is approximately 40%—i.e., you need to charge a NiCad battery 140% of its rated capacity to have it fully charged. The flooded wet battery's internal resistance can be as high as 26%, which is the charging current lost to gassing, or breaking up of water. Gel acid batteries are better at approximately 16% internal resistance and require roughly 116% of rated capacity to be fully charged. AGM batteries have an internal resistance of 2%, allowing them to be charged faster and deliver higher power.

19.10.11 LeClanche (Zinc-Carbon) Batteries

LeClanche batteries consist of a carbon anode, zinc cathode, and electrolyte solution of ammonium chloride, zinc chloride, and mercury chloride in water (called a *mix*). The nominal voltage is 1.5 V. This type of cell is quite inefficient at heavy loads, and its capacity depends considerably on the duty cycle. Less power is available when it is used without a rest period. Maximum power is produced when it is given frequent rest periods, since the voltage drops continuously under load. Shelf life is limited by the drying out of the electrolyte. A typical discharge curve is given in Fig. 19-30.

Zinc-carbon cells may be recharged for a limited number of cycles. The following information is extracted from the National Bureau of Standards Circular 965:

> The cell voltage for recharge must not be less than 1 V and should be recharged within a short time after removing from service. The ampere hours of charge should be within 120–180% of the discharge rate. The charging rate is to be low enough to distribute the recharge over 12–16 h. Cells must be put into service soon after recharging as the shelf life is poor.

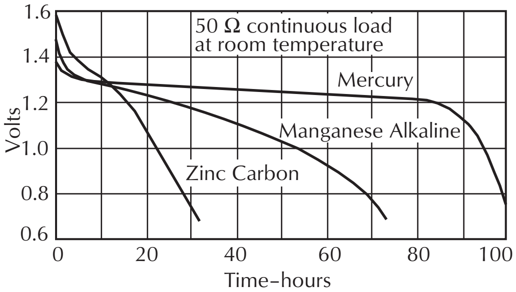


Figure 19-30. Typical discharge curves for three different types of penlight cells discharged continuously into a 50 Ω load.

19.10.12 Nickel-Cadmium Batteries

For optimum performance, many battery-operated items require a relative constant voltage supply. In most applications, nickel-cadmium cells hold an almost constant voltage throughout most of the discharge period, and the voltage level varies only slightly with different discharge rates. Nominal discharge voltage is 1.25 V at room temperature.

Nickel-cadmium cells are especially suited to high discharge or pulse currents because of their low internal resistance and maintenance of discharge voltage. They are also capable of recharge at high rates under controlled conditions. Many cells can be rapidly charged in 3–5 h without special controls, and all can be recharged at a 14 h rate.

Nickel-cadmium cells are designed to operate with a wide temperature range and can be discharged from –40°F to +140°F (–40°C to +60°C).

These cells can be continuously overcharged at recommended rates and temperature. This will not noticeably affect life unless the charge rate exceeds design limitations of the cell.

The cell construction eliminates the need to add water or electrolyte, and, under certain conditions, the cell will operate on overcharge for an indefinite period. A typical discharge curve for a cell, rated at 25 Ah and weighing approximately 2 lb, is given in Fig. 19-31.

The charge retention varies from 75% for 1 month to as low as 15% for 5 months. Storage at high temperatures will reduce high retention. Cells should be charged

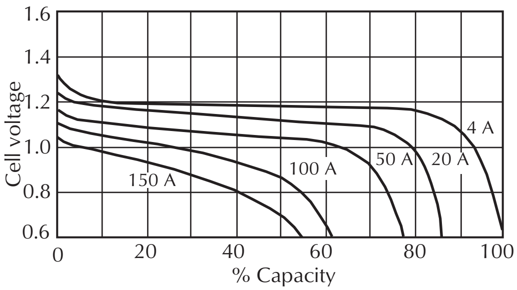


Figure 19-31. Discharge characteristic for Sonotone type 20L420 nickel-cadmium battery, rated 25 Ah.

prior to use to restore full capacity. Nickel cadmium eventually fails due to permanent or reversible cell failure. A reversible failure is usually due to shallow charge and discharge cycles and the battery appears to have lost capacity. This is often called the *memory effect*. This problem can be removed by deep discharge and a full recharge. A loss of capacity can also come from extended overcharging. If this should occur, full capacity can be restored by a discharge followed by a full recharge.

The capacity of a nickel-cadmium cell is the total amount of electrical energy that can be obtained from a fully charged cell. The capacity of a cell is expressed in ampere-hours (Ah) or milliampere-hours (mAh), which are a current-time product. The capacity value is dependent on the discharge current, the temperature of the cell during discharge, the final cutoff voltage, and the cell's general history.

The nominal capacity of the nickel-cadmium cell is that which will be obtained from a fully charged cell discharged at 68°F (20°C) for 5 h to a 1.0 V cut off. This is called the C/5 rate.

Discharges at the 20, 15, 10, and 1 h rates are called C/20, C/15, C/10, and C, respectively. Higher rates are designated as 2C, 3C, etc.

When three or more cells are series connected for higher voltages, the possibility exists that during discharge, one of the cells, which may be slightly lower in capacity than the others, will be driven to a zero potential and then into reverse. At *discharge rates* (C) in the vicinity of C/10, cells can be driven into reverse without permanently damaging the cell. Prolonged, frequent, or deep reversals should be avoided since they shorten cell life or cause it to vent. Cell voltage should never be allowed to go below –0.2 V.

Nickel-cadmium batteries may be charged using either a constant-current or constant-voltage charger. There are four major factors that determine the charge rates, which can be used on nickel-cadmium batteries. They are charge acceptance, voltage, cell pressure, and cell temperature.

No charge control is required for *charge rates* up to C/3. This allows the use of the least expensive charger design. When charging rates equal or exceed 1.0 C, the charging current must be regulated to prevent overcharge.

Table 19-7. Charging Rates for a Nickel-Cadmium Battery

Method of Charging		Charge Rate		
Name	Nickname	Current Rate	Fraction	Hour Rates
Standby	Trickle	0.01C	C/100	100 h
		0.02C	C/50	50 h
		0.03C	C/30	30 h
		0.04C	C/25	25 h
Slow	Overnight	0.05C	C/20	20 h
		0.1C	C/10	10 h
Quick	Rapid	0.2C	C/5	5 h
		0.25C	C/4	4 h
		0.3C	C/3	3 h
Fast		C.0	C	1 h
		2C.0	2C	30 min
		3C.0	3C	20 min
		4C.0	4C	15 min
		10C.0	10C	6 min

In Table 19-3, the notation that includes the letter C is used to describe current rates in terms of a fraction of the capacity rating of the battery. A comparison of cells from different manufacturers requires rationalization to a common standard for capacity rating at the same discharge rate.

In general, discharge times will be shorter than those for C rates greater than 1 and longer than those for C rates less than 1. The charge input must always be more than discharged output. For example, to ensure full recharge of a completely discharged battery, the constant-current charge time at the 10 h rate must be longer than 10 hours due to charge acceptance characteristics.

19.10.13 Alkaline-Manganese Batteries

The *alkaline-manganese battery* is gaining considerable importance in the electronic field since it is a primary battery and is rechargeable.

The polarity of this cell is reversed from the conventional zinc-carbon cell, in which the can is negative.

However, because of packaging, the outward appearance is similar to the zinc-carbon cell, with the same terminal arrangement. Although this cell has an open-circuit voltage of approximately 1.5 V, it discharges at a lower voltage than the zinc-carbon cell. Also, the discharge voltage decreases steadily but more slowly. Alkaline-manganese batteries have 50–100% more capacity than their zinc counterparts. Zinc-carbon cells yield most of their energy above 1.25 V and are virtually exhausted at 1 V, while the alkaline cell yields most of its energy below 1.25 V with a considerable portion released at less than 1 V.

If the discharge rate is limited to 40% of the nominal capacity of the cell and recharge is carried out over a period of 10–20 h, alkaline-manganese cells can be cycled 50–150 times. A typical discharge curve is shown in Fig. 19-32.

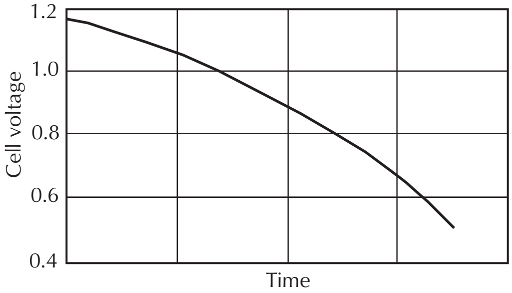


Figure 19-32. Discharge characteristics of an alkaline- manganese cell, on an arbitrary time scale.

19.10.14 Mercury Dry Cells

The *mercury dry cell* using a zinc-mercury oxide alkaline system was invented by Samuel Ruben during World War II. There are two kinds of mercury cells: one with a voltage of 1.35 V and one with 1.4 V.

The 1.35 V cell has a pure mercuric-oxide cathode. On discharge its voltage drops only slightly until close to the end of the cell life when it then drops rapidly. The 1.4 V cell has a cathode of mercuric oxide and manganese dioxide. On discharge, its voltage is not quite as well regulated as the 1.35 V cell, but it is considerably better than the manganese-alkaline or zinc-carbon cell.

Mercury cells have excellent storage stability. A typical cell will indicate a voltage of 1.3569 V, with a cell-to-cell variation of only 150 µV. Variation due to temperature is 42 µV/°F, ranging from −70°F to +70°F (−56°C to +21°C), with a slight increase of voltage with temperature. The internal resistance is approximately 0.75 Ω. Voltage loss during storage is about 360 µV per month; therefore, a single cell can be used as a reference voltage of 1.3544 V, ±0.17%. The voltage is defined under a load condition of 5% of the maximum current capacity of the cell. Normal shelf life is on the order of 3 years.

Recharging of mercury cells is not recommended because of the danger of explosion. A typical stability curve for a single cell over a period of 36 months is shown in Fig. 19-33. The drop in voltage over this period is 13 mV.

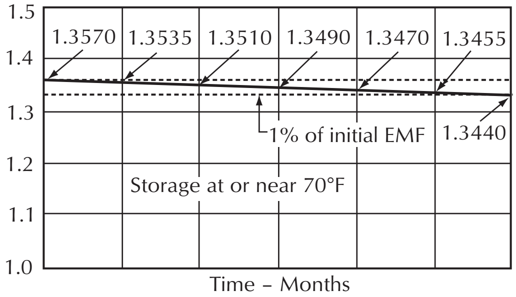


Figure 19-33. Stability curve for a single-cell mercury battery.

References

1. Kurt Mathews, Single IC, Power Factor Corrected, Off-Line Supply, *Design Note 143, Linear Technologies Design Notes*, Linear Technology Corporation.
2. Synchronous Rectification Aids Low-Voltage Power Supplies, *Maxim Acation Note 652*, Jan 31, 2001, Maxim Intergrated Products

This page intentionally left blank

Chapter 20

Amplifier Design

by Eugene Patronis, Jr.

20.1 The Necessity for Amplifiers 703
20.2 Types and Descriptions of Amplifiers 703
  20.2.1 Amplifier Transfer Function 704
  20.2.2 Feedback Theory 707
  20.2.3 Operational Amplifiers 710
  20.2.4 Active Filters Employing Operational Amplifiers 711
20.3 Power Amplifiers 716
  20.3.1 Protection Mechanisms 719
  20.3.2 High-Power Amplifiers 720
  20.3.3 High-Power Analog Amplifiers 722
  20.3.4 Switch Mode Power Supplies 724
  20.3.5 Technological Innovation 725
  20.3.6 Class D in Full Bloom 726
  20.3.7 Signal Processing in Power Amplifiers 729
  20.3.8 Computer-Controlled Power Amplifiers and Systems 730
Bibliography 731

This page intentionally left blank

20.1 The Necessity for Amplifiers

The necessity for amplification becomes apparent from an analysis of the unlikely arrangement depicted in Fig. 20-1 wherein a dynamic microphone is connected directly to a loudspeaker.

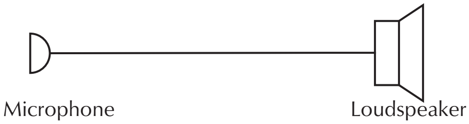


Figure 20-1. The impossible sound reinforcement system.

The microphone typically, with moderate excitation, would generate an open circuit voltage of 10 mV and possess an internal impedance of 200 Ω. The loudspeaker typically would have an impedance of 8 Ω and an efficiency of 10%. The electrical power delivered to the loudspeaker assuming that the microphone and loudspeaker impedances are predominantly resistive would be 1.8×10^{-8} W while the acoustical output of the loudspeaker would only be 1.8×10^{-9} W. Even if a matching transformer is interposed between the microphone and the loudspeaker, the improvement is hardly significant. The acoustical output in this event becomes only 1.25×10^{-8} W which is several orders of magnitude below the acoustical power requirements of most applications.

20.2 Types and Descriptions of Amplifiers

The initial description for an amplifier is based on the nature of the active elements involved such as vacuum tube, bipolar transistor, field effect transistor, integrated circuit, magnetic field, or a mixture of two or more of these technologies, in which case it is called a *hybrid*. The second descriptor is associated with the principal quantity being amplified and indirectly with the input output relationships exhibited by the amplifier.

For example, a *voltage amplifier* is excited at its input by a signal in the form of a voltage and responds by producing a related voltage at its output. In this instance, it is desirable that the input impedance of a voltage amplifier be large compared with the impedance of the signal source and that the output impedance of the amplifier be small compared with the load impedance connected at its output. As a result, the signal source impresses a maximum voltage across the amplifier's input and the amplifier subsequently produces a maximum voltage across its associated load.

A *current amplifier* is excited at its input by a signal in the form of a current and responds by producing a related current in its associated load. Current amplifiers have low input impedances and high output impedances.

A *transconductance amplifier* is excited at its input by a voltage and responds by producing a related current in its associated load. Transconductance amplifiers have high input impedances and high output impedances.

Transresistance amplifiers are excited at the input by a signal current and respond by producing a related voltage at the output. A transresistance amplifier has a low input impedance as well as a low output impedance.

Another useful amplifier descriptor describes the functional relationship, in a mathematical sense, which exists between the input and output signals. For example, in *linear amplifiers* the output signal is a linear function of the input signal whereas in *logarithmic amplifiers*, the output signal is proportional to the logarithm of the input signal. The majority of the amplifiers employed in audio are linear but a significant number of logarithmic or other special function amplifiers find use in signal processing applications.

Additional descriptions are associated with the physical location of an amplifier in the overall amplification chain. For example, a preamplifier is usually placed immediately after a transducer where the signal levels are quite low and noise characteristics are of considerable importance. Certain preamplifiers will incorporate special equalization circuitry; for instance, a phono preamplifier provides the required RIAA playback characteristic.

Preamplifiers are followed by mixing amplifiers that can combine and individually control the signals from several different sources. Although there may exist several other intermediate steps, the power amplifier is the last step.

Power amplifiers in audio work have the input-output impedance characteristics of a voltage amplifier along with the ability to deliver large amounts of electrical power. Fig. 20-2 illustrates a typical arrangement in a reinforcement chain.

The final descriptor to be discussed concerns amplifier terminal connections. Amplifiers are essentially two port devices, i.e., they are constituted with a pair of input terminals and a pair of output terminals as indicated in Fig. 20-3.

If neither of the input terminals is connected directly to ground and if both of the input terminals are electrically symmetric with respect to ground, the input is said to be balanced. If one of the input terminals is connected to ground, the input is unbalanced and is described as being single ended. Similar statements are applicable to the output pair of terminals. All possible combinations are encountered in practice. The input

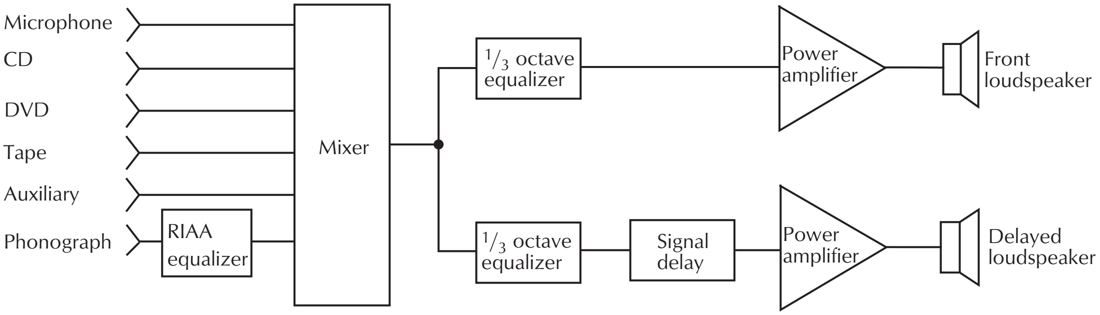


Figure 20-2. A typical reinforcement and reproduction chain.

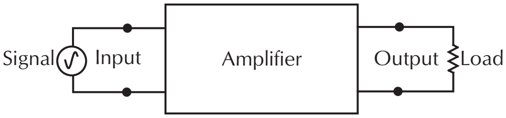


Figure 20-3. An amplifier as a two port device.

may be balanced while the output is unbalanced, the input may be unbalanced while the output is balanced, both input and output may be balanced, or both input and output may be unbalanced. The balanced configuration is preferable when dealing with long lines and low signal levels or where ground isolation is required. This preference is based on the common mode rejection properties of the balanced arrangement. For example, consider the signal leads in Fig. 20-3 to be a long twisted pair contained within an electrostatic shield with the shield grounded. The electrostatic shield offers no noise immunity from external time-varying magnetic fields. When such varying magnetic fields exist, a noise signal will be induced between each of the signal conductors and ground. The amplifier, however, amplifies the difference that appears between its input terminals and hence any common signal between the input terminals, and ground is rejected.

20.2.1 Amplifier Transfer Function

The relationship that exists in the steady state between the output signal and the input signal of a two-port device such as an amplifier or filter is called the *transfer function*. The transfer function has a magnitude and an angle with each being dependent on the steady state signal frequency. Mathematically, the transfer function is expressed concisely in the form of a complex function that has both real and imaginary parts. The magnitude of the transfer function at any particular frequency is the square root of the sum of the squares of the real and imaginary parts and physically corresponds to the ratio of the output signal amplitude to the input signal amplitude. The angle of the transfer function at any particular frequency is the angle whose tangent is the ratio of the imaginary and real parts and physically corresponds to the phase difference between the output signal and the input signal. These ideas are best expressed by a simple example. Consider a dc-coupled voltage amplifier that offers an amplification of 10 volts per volt (V/V) at dc or zero frequency and an amplification of $10/\sqrt{2}$ V/V at a frequency f_0 while having introduced a phase shift of $-\pi/4$ radian, or $-45°$. Upon denoting the transfer function by the symbol A and the independent frequency variable by the symbol f, the following statements can be made:

$$A = \left[\frac{10}{1+\left(\frac{f}{f_0}\right)^2}\right] + j\left[\frac{-10\frac{f}{f_0}}{1+\left(\frac{f}{f_0}\right)^2}\right] \tag{20-1}$$

or more compactly

$$A = G\varepsilon^{j\phi} \tag{20-2}$$

where,

G is the magnitude of the transfer function or gain function,

ε is the base of the natural logarithm,

ϕ is the angle of the transfer function or phase function, the angle whose tangent is the imaginary part divided by the real part of Eq. 20-1,

G is the square root of the sum of the squares of the real and imaginary parts of Eq. 20-1,

j is $\sqrt{-1}$.

$$G = \frac{10}{\sqrt{1+\left(\frac{f}{f_0}\right)^2}} \tag{20-3}$$

$$\phi = \tan^{-1}\left[\frac{-f}{f_0}\right] \tag{20-4}$$

An elegant form in which to express Eq. 20-1 is obtained by letting $S = j\omega$ with ω equal to $2\pi f$ and ω_0 equal to $2\pi f_0$. Eq. 20-1 after substitution and simplification becomes

$$A = \frac{10\omega_0}{S + \omega_0} \tag{20-5}$$

Eq. 20-5 is the statement of Eq. 20-1 in the language of the Laplace transform, which is really the basis for transfer function analysis. It is worthwhile at this point to note that Eqs. 20-1, 20-2, and 20-5 are alternative ways of expressing the transfer function of the simple amplifier under discussion. The form in Eq. 20-5 is that which is used most often in practice because of its simplicity.

If S were allowed to assume any possible value, whether it be real, imaginary, or complex, such that all points in a two-dimensional complex plane were accessible, there would be only one value of S in Eq. 20-5 for which the denominator would become zero and A would become infinite. That value of S is when $S = -j\omega_0$. It is said then that Eq. 20-5 has a single pole located at $S = -\omega_0$. The pole order of a transfer function is determined by the power of S appearing in the denominator. A two-pole amplifier would have an S^2, a three-pole an S^3, etc., appearing in the denominator of the transfer function. In the steady state as opposed to transient state recall that S is restricted to the values $S = j\omega$ and the only accessible points lie on the positive imaginary axis because the physical frequency values must be positive. In the steady state even though the value of S never coincides with the location of our example pole, the pole location nevertheless influences the operation of the amplifier. Changing the pole location in effect changes the value of ω_0 and hence changes the value of the transfer function at all frequencies other than zero frequency.

A further study of the Laplace transform and the inverse Laplace transform indicates that the transfer function is a description also of the device's impulse response in the complex frequency plane while the inverse Laplace transform of the transfer function is the description of the device's response to an impulse described in the time domain—i.e., it is the device's transient response to an impulse expressed as a function of time. An important consequence of this is that in order for a device to exhibit a transient response that decays with increasing time, all of the poles of the device's transfer function must have negative real parts. The amplifier under discussion satisfies this criterion with a pole at $-\omega_0$ and hence its transient response decays with time which allows the amplifier to exhibit a stable steady-state response. If this were not true, the device would not be useful as an amplifier.

The information contained in the amplifier's transfer function may be depicted in a variety of ways, the two most popular of which are the Bode and Nyquist diagrams. The Bode diagram displays Eqs. 20-3 and 20-4 in the form of a graph of $20\,\mathrm{dB}\log G$ plotted versus $\log\omega$ and a graph of ϕ plotted versus $\log\omega$. Fig. 20-4 is the Bode diagram for the amplifier of the example.

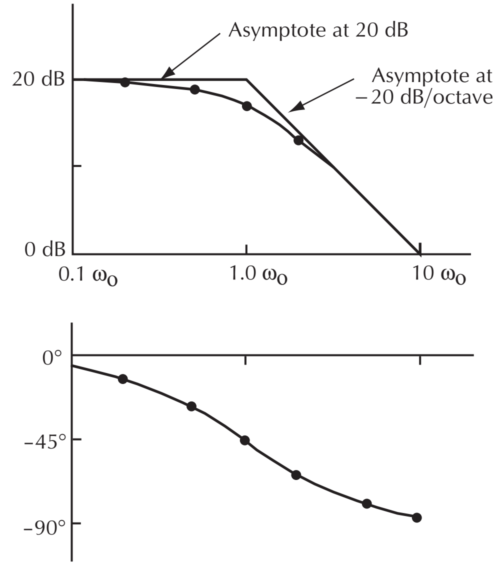


Figure 20-4. Gain and phase graphs for the example amplifier.

An examination of the Bode diagram of Fig. 20-4 leads to the conclusion that this amplifier is in essence a low-pass filter having a reference gain of 20 dB, a single pole, and a half power point at $\omega = \omega_0$. The pole order is deduced from the fact that even though the response is low pass in nature its asymptotic slope is −20 dB per decade or equivalently −6 dB per octave. A two-pole low pass would produce −12 dB per octave, a three-pole −18 dB per octave, etc., in the asymptotic slope.

This same information is displayed in a different form by means of a Nyquist diagram. A Nyquist diagram is a graph in the complex plane of Eq. 20-1 plotted under the condition that ω is allowed to take on all values from zero to infinity. Fig. 20-5 is the Nyquist diagram for the amplifier of the example.

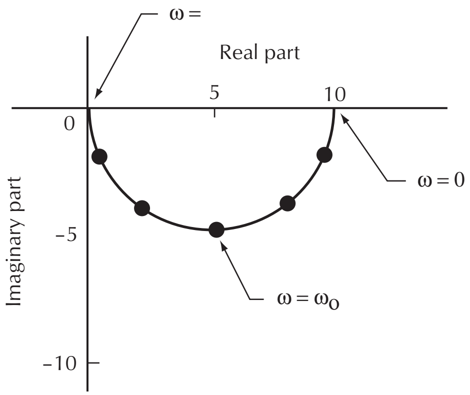


Figure 20-5. Nyquist diagram for the example amplifier.

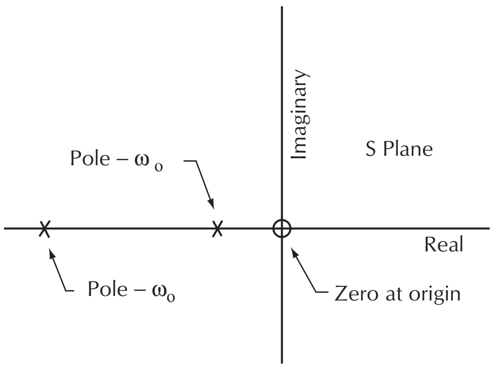


Figure 20-6. Pole-zero diagram.

A second example will serve to further explore the properties of transfer functions. Consider that the amplifier of the previous example had an input resistance of amount R. The input circuit is now to be modified by connecting a capacitor of size C in series with this input resistance to form a simple ac-coupled amplifier. Upon denoting $\omega_0' = 1/RC$ the transfer function for this new amplifier is given by

$$A = \frac{10S\omega_0}{(S+\omega_0')(S+\omega_0)} \tag{20-6}$$

Eq. 20-6 indicates that the amplifier now has two poles, the original one at $S = -\omega_0$ and a new one located at $S = -\omega_0'$. In addition, an examination of the numerator of the transfer function indicates that there is now a value of S for which the numerator becomes zero namely at $S = 0$. Values of S that make the numerator zero are called the zeros of the transfer function. The present amplifier has a single zero and a pair of poles that can be displayed in a pole-zero diagram. A pole-zero diagram is a drawing of the complex frequency plane in which the pole locations are denoted by X and the zero locations by 0. The pole-zero diagram for the ac-coupled amplifier appears as Fig. 20-6.

Fig. 20-6 is a relatively simple pole-zero diagram as the amplifier upon which it is based is simple. A few conclusions based on more general amplifiers are worth noting. Real poles may be singular while complex poles always appear as conjugate pairs. The poles for amplifiers that exhibit stable steady-state behavior may be real or complex but must have negative real parts. The zeros may appear anywhere in the S plane, but any zeros with positive real parts are associated with nonminimum phase behavior.

The Bode diagram for this example is arrived at by the following steps. First, reform Eq. 20-6.

$$A = \left(\frac{S}{S+\omega_0'}\right)\left(\frac{10\omega_0}{S+\omega_0}\right)$$

Substitute for ω_0' in terms of ω_0 by examining the pole-zero diagram, $\omega_0' = 1/4\omega_0$ therefore

$$A = \left(\frac{S}{S+\frac{\omega_0}{4}}\right)\left(\frac{10\omega_0}{S+\omega_0}\right)$$

Substitute $S = j\omega$ and find the absolute magnitude of the resulting expression to obtain the gain function as indicated here

$$\begin{aligned} G &= |A| \\ &= \frac{\omega}{\sqrt{\omega^2+\frac{\omega_0^2}{16}}} \times \frac{10\omega_0}{\sqrt{\omega^2+\omega_0^2}} \\ &= \frac{\frac{\omega}{\omega_0}}{\sqrt{\frac{\omega^2}{\omega_0^2}+\frac{1}{16}}} \times \frac{10}{\sqrt{\frac{\omega^2}{\omega_0^2}+1}} \end{aligned}$$

Make a graph of $20\ \text{dB}\log G$ versus $\log(\omega/\omega_0)$. This graph appears in Fig. 20-7. Next determine the phase function ϕ.

$$\begin{aligned} A &= \frac{S}{S+\dfrac{\omega_0}{4}} \times \frac{10\omega_0}{S+\omega_0} \\ &= \frac{j\omega}{j\omega+\dfrac{\omega_0}{4}} \times \frac{10\omega_0}{j\omega+\omega_0} \end{aligned}$$

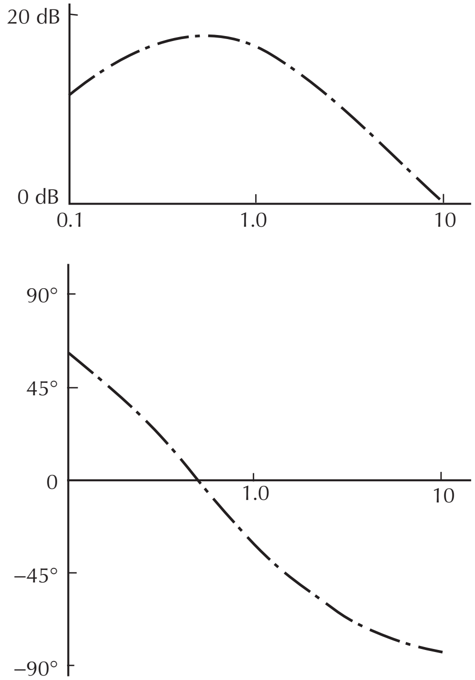


Figure 20-7. Bode diagram for simple ac amplifier.

The angle of the first factor is

$$\frac{\pi}{2} - \tan^{-1}\left(\frac{4\omega}{\omega_0}\right)$$

while the angle of the second factor is

$$-\tan^{-1}\left(\frac{\omega}{\omega_0}\right)$$

therefore, the total phase shift is

$$\phi = \frac{\pi}{2} - \tan^{-1}\left(\frac{4\omega}{\omega_0}\right) - \tan^{-1}\left(\frac{\omega}{\omega_0}\right)$$

Make a graph of ϕ versus $\log(\omega/\omega_0)$ in order to complete the Bode diagram. This graph also appears in Fig. 20-7.

20.2.2 Feedback Theory

Fig. 20-8 represents a generalized feedback loop based on a voltage amplifier. In the absence of feedback, with the loop open, the amplifier has a transfer function A. The feedback path has a transfer function B, the input signal from the outside world is V_{in} and the signal supplied as an output is V_0. When the loop is closed, the input signal is combined with the feedback signal in the indicated junction to form an error signal V_e. The process that occurs in the junction may be either addition or subtraction, depending on the nature of A, B, and the type of feedback (positive or negative) desired. In any event the signal actually supplied to the amplifier when the loop is closed is V_e.

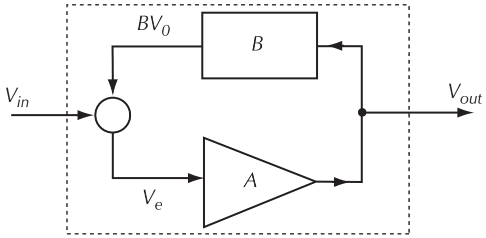


Figure 20-8. Generalized feedback loop.

The system contained within the dotted enclosure of Fig. 20-8 has a rather different transfer function from that of the amplifier operated under open loop conditions. The closed loop transfer function denoted by A' is derived as follows:

$$V_e = V_{in} + BV_0 \tag{20-7}$$

$$\begin{aligned} V_0 &= AV_e \\ &= AV_{in} + ABV_0 \end{aligned} \tag{20-8}$$

$$\begin{aligned} A' &= \frac{V_0}{V_{in}} \\ &= \frac{A}{1-AB}. \end{aligned} \tag{20-9}$$

A and B, in general, are complex functions of the steady-state frequency of operation. The absolute magnitude of the denominator of Eq. 20-9 is called the *gain reduction factor*. The feedback is called negative when

$$|1-AB| > 1 \tag{20-10}$$

and is positive when

$$|1 - AB| < 1 \, . \tag{20-11}$$

The nature of the feedback is best explored by studying the quantity AB as a function of the frequency as displayed in a Nyquist diagram. The quantity AB is called the *loop gain* of the hypothetical amplifier. Included in the diagram is a circle of unit radius centered on the point $1, j0$.

The perimeter of the unit circle divides the plane into two regions. For all of the points outside of the circle $|1 - AB| > 1$, the feedback is negative, whereas for any point on the curve within the unit circle, $|1 - AB| < 1$, the feedback is positive. The hypothetical amplifier for which Fig. 20-9 was drawn thus has negative feedback at low frequencies but exhibits positive feedback over a range of high frequencies. Note that AB is negative, real, and has its maximum absolute value at $\omega = 0$. This is characteristic of a dc-coupled amplifier having a loop gain transfer function that has poles but no zeros. Furthermore, AB resides in the second quadrant until ω exceeds ω_1 for $\omega_1 < \omega < \omega_2$, AB is in the first quadrant but the feedback is still negative. Whenever the frequency is such that $\omega > \omega_2$, AB falls within the unit circle and the feedback becomes positive. This region must be handled with extreme care.

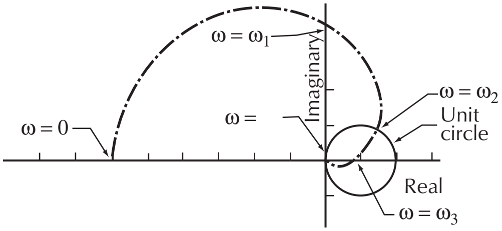


Figure 20-9. Nyquist diagram for loop gain.

As will presently be discussed in detail, negative feedback is a stabilizing influence on amplifier performance but positive feedback is a destabilizing influence and can, in fact, lead to an uncontrolled oscillatory condition. The final conclusion to be drawn from Fig. 20-9 is that for $\omega > \omega_3$, AB is in the fourth quadrant and finally approaches zero as the operating frequency becomes very large. As ω is allowed to vary from zero to infinity, the angle associated with AB undergoes a change of −270°, which is characteristic of a transfer function that, in the absence of zeros, possesses three poles. If AB for the hypothetical amplifier had possessed just a single pole, the entire Nyquist diagram would have been restricted to the second quadrant and the feedback would have been negative for all frequencies. On the other hand, if AB had possessed just two poles, the Nyquist diagram would enter the first quadrant at high frequencies but would approach zero without ever crossing the real axis. The feedback would be positive at high frequencies but not to an excessive degree.

The critical point to be avoided for stable operation is the point $1, j0$ on the positive real axis. If the Nyquist curve passes through this point under any condition, the loop gain becomes one with an angle of zero. As a consequence, $|1-AB|$ becomes zero and A' becomes infinite. Physically this implies that the amplifier will produce an output even in the absence of an input signal from the outside world. That is, what was intended to be an amplifier has become an oscillator. The type of Nyquist diagram that is to be avoided is one that encircles the critical point such as displayed in Fig. 20-10. Fig. 20-10 was obtained by a modification of the amplifier described by Fig. 20-9. This modification amounted to changing AB at zero frequency from its former value of −6 to a new value of −10 with all other factors remaining the same.

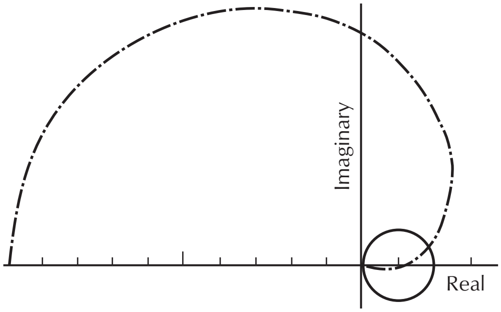


Figure 20-10. Nyquist curve for an unstable amplifier.

Unlike Fig. 20-9, Fig. 20-10 is the Nyquist curve for an unstable amplifier. The curve does not pass through the critical point but it does encircle the critical point. Consider for the moment that the amplifier is initially off. Under such a circumstance, A is zero and, consequently, AB is also zero. Under these conditions the Nyquist curve is collapsed into a single point at the origin. Following turn-on, there is a period of time in which A and AB are growing toward their final values. In this interval, the Nyquist curve is in effect growing outward from the origin. At some instant during this growth period, the Nyquist curve will intersect the critical point $1, j0$ and the amplifier will break into oscillation. Precaution must be taken, therefore, when dealing with feedback loops in which the loop gain transfer function has three or more poles.

It was mentioned earlier that negative feedback can be a stabilizing influence on amplifier operation. A negative feedback loop is in essence a type of quality control wherein the system output is compared with what is desired for it to be. Any difference as a result of this comparison is injected back into the system in such a way as to force a correction of system behavior. A highly simplified example is as follows. Consider a dc voltage-amplifier for which it is desired that the voltage gain be –10; that is, an output voltage ten times as large as the input signal but with the opposite polarity. One might proceed on good faith and employ the latest electronic design techniques, consult manufacturer's specifications on the best available active devices, design, and finally construct an amplifier that, according to the best available information, possesses an open loop transfer function at low frequencies of –10. In fact, in order to be on the safe side, one may follow the same procedure yielding a value of –20 and precede the device by an adjustable attenuator set at an absolute value of ½ or whatever is required to obtain an overall transfer function of –10 when the system is first tested. Unfortunately, the active devices employed are at the mercy of the operating voltages supplied to them (line voltage variations, etc.), ambient temperature variations, age, and weather elements in general. To a lesser degree, the same may be said of the passive elements involved. A, the open loop transfer function may possess a nominal value of –10 but it is constantly changing from moment to moment being at times larger and at other instants smaller than the intended value. There exists nothing in the system to monitor its overall operation. Alternatively, one might, following the same procedures outlined before, design an amplifier having an open loop transfer function whose nominal value is –100 and enclose this with a negative feedback loop to obtain a nominal closed loop transfer function, A', of –10. Mathematically,

$$A' = \frac{A}{1 - AB}$$

By substituting nominal values, one can solve for B:

$$\begin{aligned} -10 &= \frac{-100}{1 + 100B} \\ 1 + 100B &= 10 \\ 100B &= 9 \\ B &= \frac{9}{100} \end{aligned}$$

B is found to require the properties of a simple attenuator or voltage divider. The next step would be to construct this divider from precision resistors possessing very small voltage and temperature coefficients of resistance. The feedback loop is then closed, making use of this stable attenuator. The resulting system has a nominal closed loop transfer function, $A' = -10$, a nominal loop gain, $AB = -9$, and a nominal gain reduction factor of 10. What has been accomplished? Suppose that the original open loop amplifier whose nominal transfer function was –10 had variations or changes in A that were about ±20% and the new amplifier that was constructed employing the same technology has similar variations under open loop conditions. Now compare the ratio of the variations to the nominal values with and without feedback; that is, $\Delta A/A$ is to be compared with $\Delta A'/A'$. Knowing that $A' = A/(1 - AB)$ and by employing the techniques of differential calculus one finds that

$$\Delta A' = \frac{\Delta A}{(1 - AB)^2}$$

Consequently,

$$\begin{aligned} \frac{\Delta A'}{A'} &= \frac{\dfrac{\Delta A}{(1 - AB)^2}}{\dfrac{A}{1 - AB}} \\ &= \frac{\Delta A}{A} \times \frac{1}{1 - AB} \\ &= \pm 20\% \times \frac{1}{10} \\ &= \pm 2\% \end{aligned}$$

The application of negative feedback has produced a system that has a nominal transfer function of –10 with a variation of ±2%; whereas before, in the absence of feedback, there existed a system having a nominal transfer function of –10 with a variation of ±20% under the same conditions. The price paid for this improvement amounted to trading off a higher open loop gain for the sake of a more stable value of gain.

Negative feedback affects many amplifier properties other than gain stability. Negative feedback increases amplifier bandwidth, reduces most but not all forms of distortion, modifies amplifier input and output impedances, and can be beneficially employed in shaping frequency response characteristics. Examples of these features are given in the next section.

Negative feedback is not, however, a panacea. It can not turn a bad amplifier into a good one. It may make a good amplifier into a better one. It should always be remembered that the derivations and conclusions obtained above are based on linear or nearly linear operating conditions of the active devices. Negative feedback loops lose control under clipping conditions and recovery from such conditions may be poorer with negative feedback than without it.

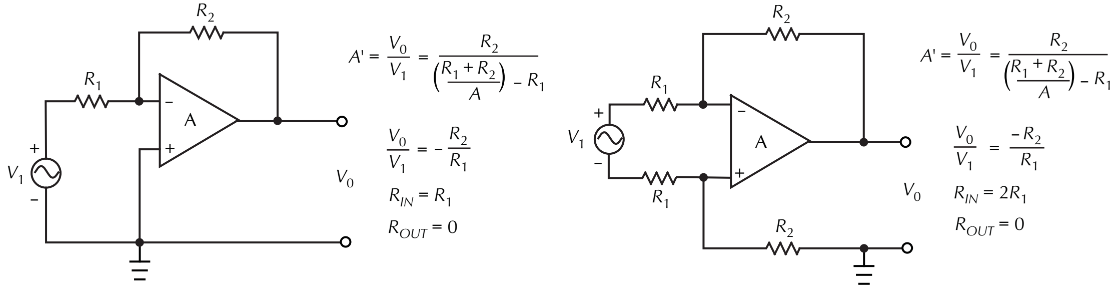
20.2.3 Operational Amplifiers

Operational amplifiers derive the name as a result of their first employment in analog computing systems. In this role, with suitable feedback, they were employed to accomplish the mathematical operations of addition, subtraction, integration, and differentiation. In their current form of integrated circuits, operational amplifiers have become the fundamental building blocks of electronic analog circuits with notable uses in power supply regulation, voltage and current amplification, and active filters, as well as other forms of signal processors.

Operational amplifiers are dc-coupled voltage amplifiers possessing, under open loop conditions, very high gain, wide bandwidth, high input impedance, low output impedance, balanced or difference inputs accompanied usually by a single-ended output, and provisions for accomplishing a dc voltage balance at the output.

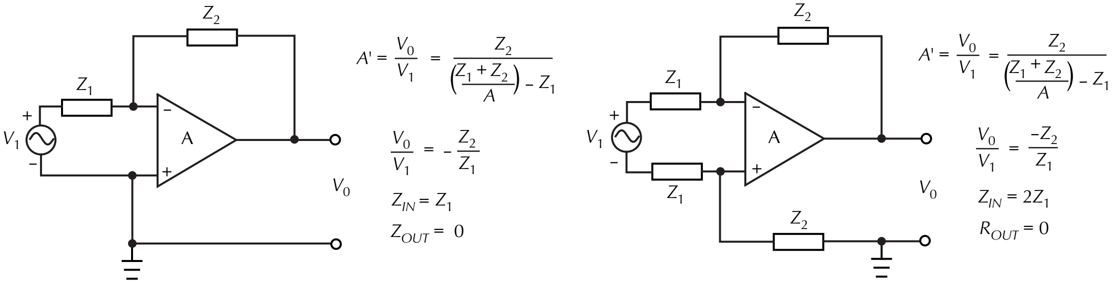
Fig. 20-11 displays the configurations commonly employed for operational amplifiers where signal inversion (polarity change) is required or desirable. In each instance the open loop transfer function, *A*, is negative and real at low frequencies.

In Fig. 20-11A through 20-11D expressions are given for the respective closed loop transfer functions *A'*. These expressions are valid without correction provided that the input impedance of the operational amplifier under open loop conditions is much larger than the impedances used in structuring the loop and that the output impedance of the operational amplifier


R_2
R_1
A
V_1
V_0
$A' = \frac{V_0}{V_1} = \frac{R_2}{\left(\frac{R_1+R_2}{A}\right) - R_1}$
$\frac{V_0}{V_1} = -\frac{R_2}{R_1}$
$R_{IN} = R_1$
$R_{OUT} = 0$
R_2
R_1
R_1
R_2
A
V_1
V_0
$A' = \frac{V_0}{V_1} = \frac{R_2}{\left(\frac{R_1+R_2}{A}\right) - R_1}$
$\frac{V_0}{V_1} = \frac{-R_2}{R_1}$
$R_{IN} = 2R_1$
$R_{OUT} = 0$

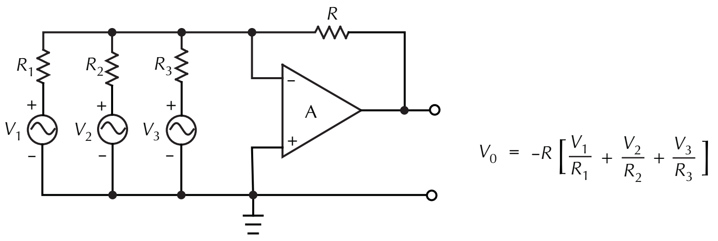
A. Unbalanced input.

B. Balanced input.


Z_2
Z_1
A
V_1
V_0
$A' = \frac{V_0}{V_1} = \frac{Z_2}{\left(\frac{Z_1+Z_2}{A}\right) - Z_1}$
$\frac{V_0}{V_1} = -\frac{Z_2}{Z_1}$
$Z_{IN} = Z_1$
$Z_{OUT} = 0$
Z_2
Z_1
Z_1
Z_2
A
V_1
V_0
$A' = \frac{V_0}{V_1} = \frac{Z_2}{\left(\frac{Z_1+Z_2}{A}\right) - Z_1}$
$\frac{V_0}{V_1} = \frac{-Z_2}{Z_1}$
$Z_{IN} = 2Z_1$
$R_{OUT} = 0$

C. Unbalanced but more versatile than A.

D. Balanced but more versatile than B.


R
R_1
R_2
R_3
V_1
V_2
V_3
A
$V_0 = -R\left[\frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}\right]$

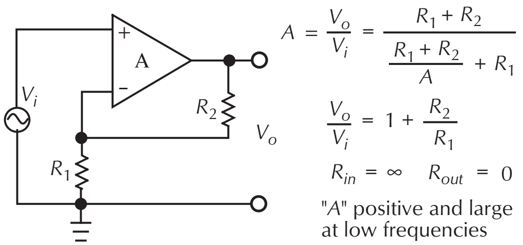
E. Combining amplifier.

Figure 20-11. Inverting operational amplifier circuits.

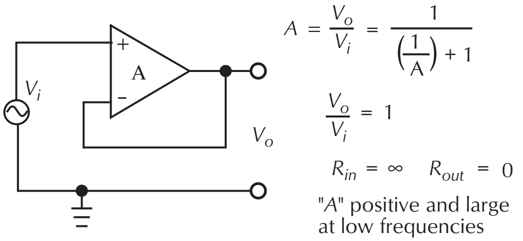
under open loop conditions is much smaller than any of the impedances used in structuring the loop. These requirements are easily met in practice as input impedances of commercial devices range upward from several megohms while the output impedances range downward from several tens of ohms throughout the audio frequency range. The approximate values given for V_o/V_i are valid if, in addition to the requirements stated above, the magnitude of A is large throughout the frequency range being employed. Fig. 20-11A is an inverting voltage amplifier having an unbalanced input as well as output. Fig. 20-11B is an inverting voltage amplifier with a balanced input. Fig. 20-11C (unbalanced) and Fig. 20-11D (balanced) are examples of more versatile configurations. The impedances Z_1 and Z_2 can be any two terminal configurations of impedance elements. These circuits find applications as low-pass filters or integrators, high-pass filters or differentiators, phase compensators, shelving filters, and tone controls among a myriad of other possibilities. Fig. 20-11E is a combining amplifier that combines or adds signals from several sources with different weighting or gain factors for each signal.

Fig. 20-12A is an example of a noninverting voltage amplifier and Fig. 20-12B is a noninverting unity gain voltage follower that is often employed as a buffer because of its extremely high input impedance and exceptionally low output impedance. In each instance the open loop transfer function, A, is positive and real at low frequencies.

Most of the wideband low noise operational amplifiers currently available for audio applications are internally structured so as to exhibit dominant pole characteristics. This means that the open loop transfer function of such an amplifier exhibits the behavior of a single-pole amplifier over the frequency range for which it is useful. Such an amplifier is easily employed in the majority of feedback arrangements without fear of violating the conditions necessary for stability. Fig. 20-13 is a Bode diagram typical of such amplifiers, both when operated open loop as well as when operated with a closed loop noninverting voltage gain of 20 dB.


+
A
–
V_i
R_2
V_o
R_1
$A = \frac{V_o}{V_i} = \frac{R_1 + R_2}{\frac{R_1 + R_2}{A} + R_1}$
$\frac{V_o}{V_i} = 1 + \frac{R_2}{R_1}$
$R_{in} = \infty \quad R_{out} = 0$
"A" positive and large at low frequencies

A. Noninverting voltage amplifier.


+
A
–
V_i
V_o
$A = \frac{V_o}{V_i} = \frac{1}{\left(\frac{1}{A}\right) + 1}$
$\frac{V_o}{V_i} = 1$
$R_{in} = \infty \quad R_{out} = 0$
"A" positive and large at low frequencies

B. Noninverting unity gain voltage amplifier.

Figure 20-12. Noninverting voltage amplifiers.

An examination of Fig. 20-13 reveals that under open loop conditions this amplifier exhibits a gain of 90 dB or $\sqrt{10} \times 10^4$ V/V at dc with the gain being down by 3 dB at a frequency of $\sqrt{10} \times 10^2$ Hz attended by a phase shift of –45°. The bandwidth of this amplifier is then $\sqrt{10} \times 10^2$ Hz and the product of the gain at dc with the bandwidth or the gain bandwidth product is 10^7 HzV/V. The loop is closed in this example by requiring that R_2 in Fig. 20-12A be nine times the value of R_1. The second set of curves in Fig. 20-13 describe the performance under this closed loop condition. The curves reveal that the gain at dc is now 20 dB or 10 V/V and that the bandwidth has now become 10^6 Hz. The gain bandwidth product is still 10^7 HzV/V. The bandwidth has been increased by exactly the same factor that the gain was reduced. This behavior is characteristic of dominant pole amplifiers. The application of feedback has yielded another important benefit. The open loop amplifier not only had a nonflat amplitude response throughout most of the audio spectrum, it suffered from phase or group delay distortion above a few hertz as well. The amplifier with feedback has a linear phase behavior from dc to beyond 10^4 Hz and hence does not introduce any group delay distortion in this frequency range.

20.2.4 Active Filters Employing Operational Amplifiers

Filter technology has a long time-honored history that actually predates electronics by several decades. In fact, if Lord Kelvin (William Thomson) had not discovered the physical and mathematical properties of so-called wave filters in the middle of the 1800s, submarine telegraphic cable communications and later long distance telephone communications would have been delayed until well into the 20th century.

In spite of the voluminous literature and interest in this subject, what will be touched on here are just a few of the filter types that have proven to be of paramount importance in modern audio applications and more particularly those that are readily implemented by means of active circuitry. Even though the emphasis will be on active circuitry, a cursory examination of some simple passive structures is of value.

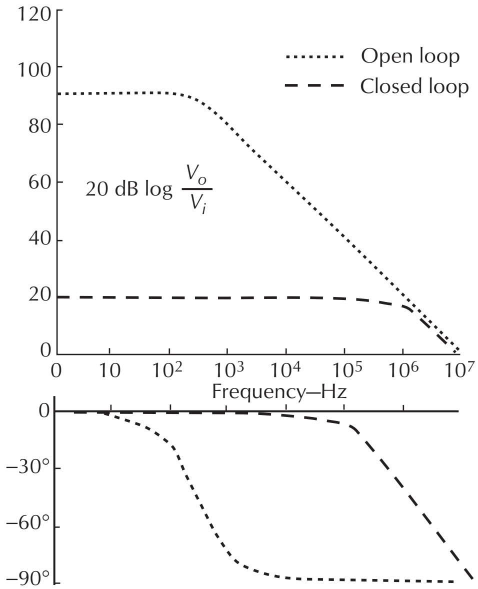
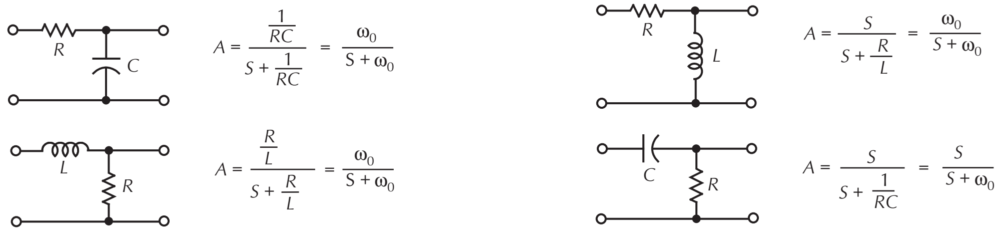


Figure 20-13. Bode diagram for a dominant pole amplifier.

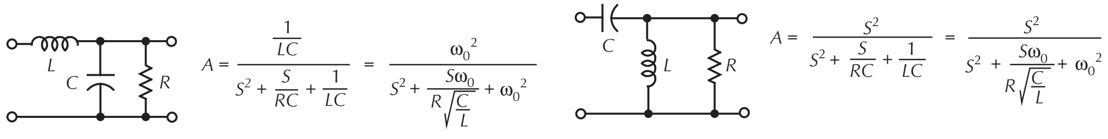
Fig. 20-14 displays a few passive filter structures along with their associated transfer functions. Note that in each instance the filter transfer functions involve a polynomial with S appearing in the denominator. The characteristics of the various filters are associated with the structure of these polynomials. With the possible exception of antialiasing use, the filters employed in audio work are restricted to pole orders of three or less in order to maintain good transient response. A pole order of three corresponds to an asymptotic slope rate of −18 dB per octave in the filter stop band. The most popular polynomials for audio applications are the Butterworth, with maximally flat amplitude response, and the Bessel, with linear phase response (maximally flat group delay). The Butterworth polynomials through third order are:

1. $S + \omega_0$
2. $S^2 + \sqrt{2}S\omega_0 + \omega_0^2$
3. $S^3 + 2S^2\omega_0 + 2S\omega_0^2 + \omega_0^3$


R
C
$A = \frac{\frac{1}{RC}}{S+\frac{1}{RC}} = \frac{\omega_0}{S+\omega_0}$
L
R
$A = \frac{\frac{R}{L}}{S+\frac{R}{L}} = \frac{\omega_0}{S+\omega_0}$
R
L
$A = \frac{S}{S+\frac{R}{L}} = \frac{\omega_0}{S+\omega_0}$
C
R
$A = \frac{S}{S+\frac{1}{RC}} = \frac{S}{S+\omega_0}$

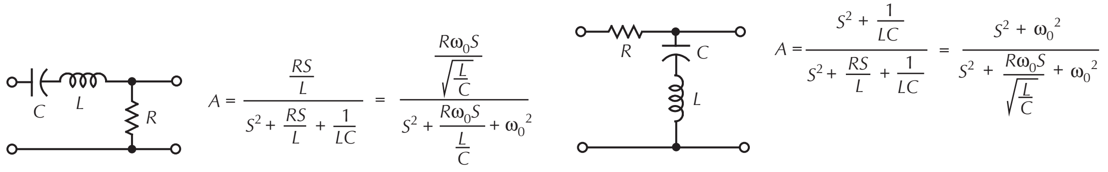
A. Single-pole low-pass.

B. Single-pole high-pass.


L
C
R
$A = \frac{\frac{1}{LC}}{S^2+\frac{S}{RC}+\frac{1}{LC}} = \frac{\omega_0^2}{S^2+\frac{S\omega_0}{R\sqrt{\frac{C}{L}}}+\omega_0^2}$
C
L
R
$A = \frac{S^2}{S^2+\frac{S}{RC}+\frac{1}{LC}} = \frac{S^2}{S^2+\frac{S\omega_0}{R\sqrt{\frac{C}{L}}}+\omega_0^2}$

C. Two-pole low-pass.

D. Two-pole high-pass.


C
L
R
$A = \frac{\frac{RS}{L}}{S^2+\frac{RS}{L}+\frac{1}{LC}} = \frac{\frac{R\omega_0 S}{\sqrt{\frac{L}{C}}}}{S^2+\frac{R\omega_0 S}{\frac{L}{C}}+\omega_0^2}$
R
C
L
$A = \frac{S^2+\frac{1}{LC}}{S^2+\frac{RS}{L}+\frac{1}{LC}} = \frac{S^2+\omega_0^2}{S^2+\frac{R\omega_0 S}{\sqrt{\frac{L}{C}}}+\omega_0^2}$

E. Two-pole bandpass.

F. Two-pole band reject.

Figure 20-14. Passive filter structures.

These are often written in a normalized form such as

1. $\frac{S}{\omega_0} + 1$
2. $\frac{S^2}{\omega_0} + \frac{\sqrt{2}S}{\omega_0} + 1$
3. $\frac{S^3}{\omega_0^{\ 3}} + \frac{2S^2}{\omega_0^{\ 2}} + \frac{2S}{\omega_0} + 1$

In the Butterworth polynomials, $\omega_0 = 2\pi f_0$, where f_0 is the frequency at which the response is 3 dB down. The Butterworth polynomials yield excellent amplitude response characteristics while their phase and group delay characteristics are far from being ideal. Their use in constant resistance crossover networks is almost universal.

The Bessel polynomials in normalized form are

1. $\frac{S}{\omega_0} + 1$
2. $\frac{S^2}{3\omega_0^{\ 2}} + \frac{S}{\omega_0} + 1$
3. $\frac{S^3}{15\omega_0^{\ 3}} + \frac{2S^2}{5\omega_0^{\ 2}} + \frac{2S}{\omega_0} + 1$

Here, ω_0 has the significance that the group delay at zero frequency is just the reciprocal of ω_0. The group delay for any filter at any frequency is given by the negative of the first derivative of phase response with respect to ω:

$$t_g = \frac{-d\phi}{d\omega} \tag{20-12}$$

For a system not to introduce any phase distortion, it is necessary that ϕ be either independent of frequency or of the form

$$\phi = -k\omega + \text{constant} \tag{20-13}$$

In the first instance $t_g = 0$ and for Eq. 20-13 $t_g = k$, where k is a constant. Bessel filters are nearly ideal in this respect as their group delays are constant or nearly so throughout their passbands. Unfortunately, the amplitude response of Bessel filters for orders higher than one, though without ripples, is not as flat as the corresponding Butterworth filter. The first-order Bessel and Butterworth filters are identical.

Operational amplifiers make significant contributions in the area of active filter implementation. The following examples, though by no means exhaustive, will serve as an introduction to this important subject.

The circuit of Fig. 20-15 simulates a physical inductor. A physical inductor at low frequencies, where interturn capacitance is not of importance, can be thought of as a pure resistance in series with a pure self-inductance. As such, a physical inductor has an impedance Z that has both a real and an imaginary part.

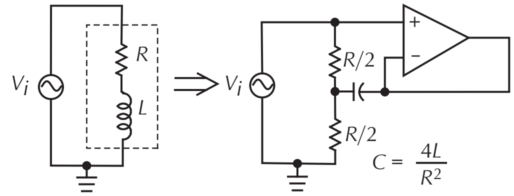


Figure 20-15. The resistor, capacitor, operational amplifier combination presents the signal source with the same impedance at all frequencies as does the physical inductor in the dotted enclosure. The two circuits are equivalent.

A physical inductor also has a quality factor or Q. These properties are summarized by the following equations:

$$Z = R + j\omega L \tag{20-14}$$

$$Q = \frac{\omega L}{R} \tag{20-15}$$

Third-order or higher filters are readily obtained by cascading two or more sections of the examples displayed in Fig. 20-16. The transfer functions of the various filters appear in Table 20-1.

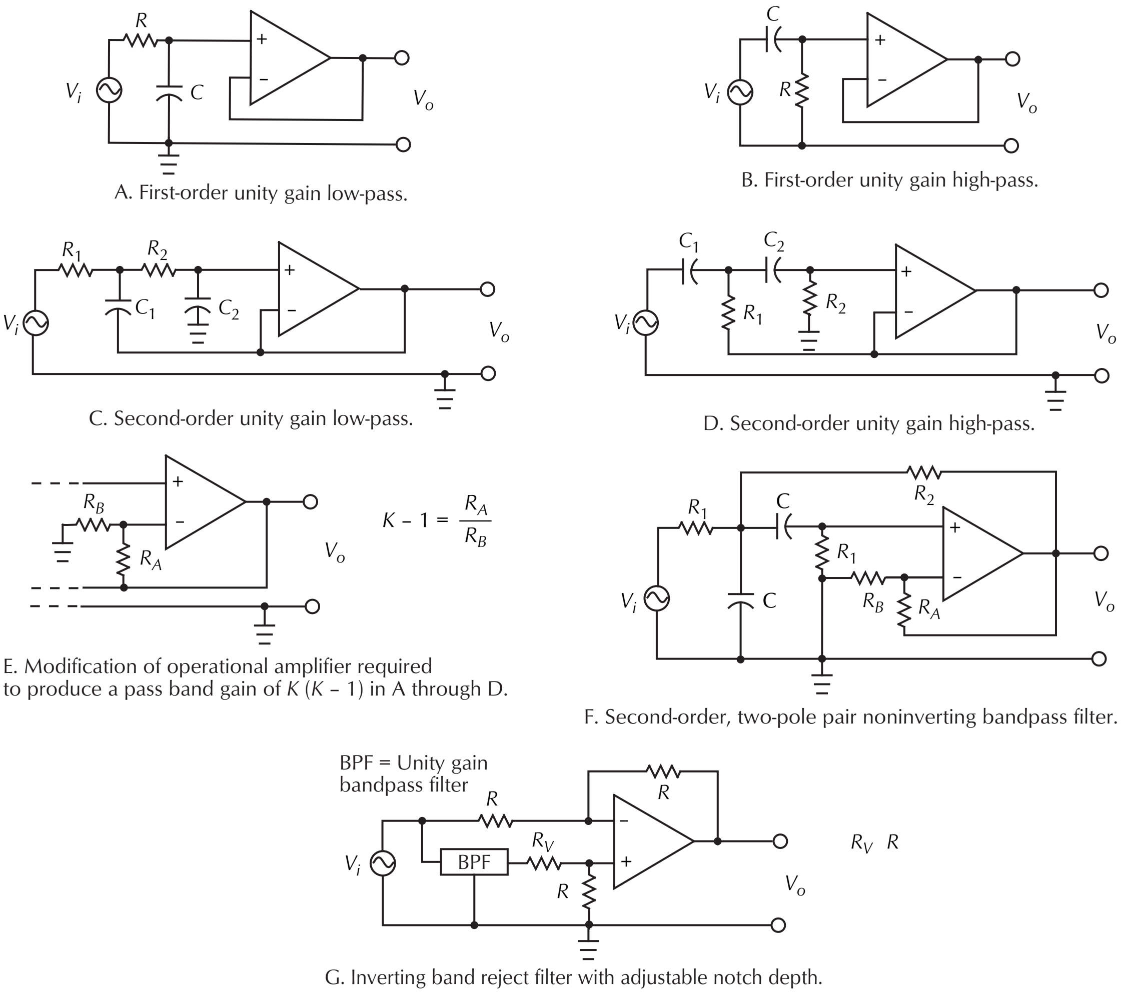
This discussion of active filters employing operational amplifiers will now be concluded by exploring two design examples.

Example 1. Third-order Butterworth low pass with a corner frequency f_0 of 500 Hz and unity gain.

This filter can be implemented by cascading a first-order section followed by a second-order section. The required overall transfer function is

$$\frac{V_0}{V_{in}} = \frac{\omega_0}{S + \omega_0} \times \frac{\omega_0^{\ 2}}{S^2 + S\omega_0 + \omega_0^{\ 2}} \tag{20-16}$$

Taking Fig. 20-16A for the first-order section along with its transfer function leads to the identification



A. First-order unity gain low-pass.

B. First-order unity gain high-pass.

C. Second-order unity gain low-pass.

D. Second-order unity gain high-pass.

E. Modification of operational amplifier required to produce a pass band gain of *K* (*K* – 1) in A through D.

F. Second-order, two-pole pair noninverting bandpass filter.

G. Inverting band reject filter with adjustable notch depth.

Figure 20-16. Active high-pass, low-pass, and band-pass filter.

$$\frac{\omega_0}{S+\omega_0} = \frac{\frac{1}{RC}}{S+\frac{1}{RC}} \tag{20-17}$$

from which it is found that

$$\begin{aligned} \omega_0 &= \frac{1}{RC} \\ &= 2\pi f_0 \\ &= 2\pi 500 \text{ Hz.} \end{aligned} \tag{20-18}$$

By choosing for C a value of 0.02 μF, Eq. 20-18 yields a value for R of 15.9 kΩ.

Taking Fig. 20-16C for the second-order section along with its transfer function leads to the identification

$$\frac{{\omega_0}^2}{S^2+S\omega_0+{\omega_0}^2} = \frac{\frac{1}{R_1R_2C_1C_2}}{S^2+\left(\frac{R_1C_2+R_2C_2}{R_1R_2C_1C_2}\right)S+\frac{1}{R_1R_2C_1C_2}}$$

$$\begin{aligned} \omega_0 &= \frac{1}{\sqrt{R_1R_2C_1C_2}} \\ &= 2\pi 500 \text{ Hz} \end{aligned} \tag{20-19}$$

and

$$\frac{1}{\sqrt{R_1 R_2 C_1 C_2}} = \left(\frac{R_1 C_2 + R_2 C_2}{R_1 R_2 C_1 C_2}\right) \tag{20-20}$$

Upon choosing $R_1 = R_2$, Eq. 20-20 dictates that $C_1 = 4C_2$. If C_1 is chosen to be 0.02 µF, then C_2 becomes 0.005 µF and Eq. 20-19 then requires that R_1 be 10 kΩ. The reasonableness of these values allows the design to be concluded with the circuit of Fig. 20-17.

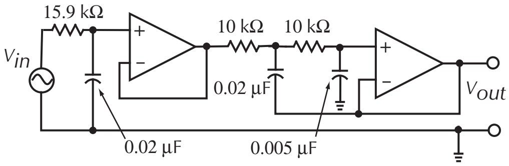


Figure 20-17. Third-octave unity gain Butterworth low-pass filter with f_o = 500 Hz.

Example 2. Second-order Bessel low pass with a zero frequency group delay of 500 µs.

The required transfer function is

$$\frac{V_0}{V_{in}} = \frac{3\omega_0{}^2}{S^2 + 3\omega_0 S + 3\omega_0{}^2} \tag{20-21}$$

with ω_0 = 1/500 µs. Taking Fig. 20-16C along with its transfer function leads to the identification

$$\frac{3\omega_0{}^2}{S^2 + 3\omega_0 S + 3\omega_0{}^2} = \frac{\dfrac{1}{R_1 R_2 C_1 C_2}}{S^2 + \left(\dfrac{R_1 C_2 + R_2 C_2}{R_1 R_2 C_1 C_2}\right) S + \dfrac{1}{R_1 R_2 C_1 C_2}} \tag{20-22}$$

from which it is found that

$$\omega_0 = \frac{1}{\sqrt{3 R_1 R_2 C_1 C_2}} \tag{20-23}$$

and

$$\frac{3}{\sqrt{3 R_1 R_2 C_1 C_2}} = \frac{R_1 C_2 + R_2 C_2}{R_1 R_2 C_1 C_2} \tag{20-24}$$

Upon choosing $R_1 = R_2$, Eq. 20-24 requires that

Table 20-1. Transfer Functions of Various Circuits of Figure 20-16

Figure	K	Transfer function
20-16A	= 1	$\dfrac{V_o}{V_i} = \dfrac{\dfrac{1}{RC}}{S + \dfrac{1}{RC}}$
20-16B	= 1	$\dfrac{V_o}{V_i} = \dfrac{S}{S + \dfrac{1}{RC}}$
20-16C	= 1	$\dfrac{V_o}{V_i} = \dfrac{\dfrac{1}{R_1 R_2 C_1 C_2}}{S^2 + \left(\dfrac{R_1 C_2 + R_2 C_2}{R_1 R_2 C_1 C_2}\right) S + \dfrac{1}{R_1 R_2 C_1 C_2}}$
20-16D	= 1	$\dfrac{V_o}{V_i} = \dfrac{S^2}{S^2 + \left(\dfrac{R_1 C_2 + R_1 C_1}{R_1 R_2 C_1 C_2}\right) S + \dfrac{1}{R_1 R_2 C_1 C_2}}$
20-16A	≥ 1	$\dfrac{V_o}{V_i} = \dfrac{K\left(\dfrac{1}{RC}\right)}{S + \dfrac{1}{RC}}$
20-16B	≥ 1	$\dfrac{V_o}{V_i} = \dfrac{KS}{S + \dfrac{1}{RC}}$
20-16C	≥ 1	$\dfrac{V_o}{V_i} = \dfrac{K\left(\dfrac{1}{R_1 R_2 C_1 C_2}\right)}{S^2 + \left(\dfrac{R_1 C_2 + R_2 C_2 + (1-K) R_1 C_1}{R_1 R_2 C_1 C_2}\right) S + \dfrac{1}{R_1 R_2 C_1 C_2}}$
20-16D	≥ 1	$\dfrac{V_o}{V_i} = \dfrac{KS^2}{S^2 + \left(\dfrac{R_1 C_2 + R_1 C_1 + (1-K) R_2 C_2}{R_1 R_2 C_1 C_2}\right) S + \dfrac{1}{R_1 R_2 C_1 C_2}}$
20-16F	≥ 1	$\dfrac{V_o}{V_i} = \dfrac{\dfrac{KS}{R_1 C}}{S^2 + \left(\dfrac{3 R_1 R_2 C + (1-K) R_1^2 C}{R_1^2 R_2 C^2}\right) S + \dfrac{R_1 + R_2}{R_1^2 R_2 C^2}}$ or $\dfrac{V_o}{V_i} = \dfrac{\left(\dfrac{A_o}{Q}\right) \omega_o S}{S^2 + \left(\dfrac{\omega_o}{Q}\right) S + \omega_o^2}$

Table 20-1. Transfer Functions of Various Circuits of Figure 20-16 (Continued)

Figure	K	Transfer function
20-16	≥ 1	where A_0 is the gain at pass band center $A_o = \frac{R_2}{\frac{1}{K}(3R_2 + R_1 - KR_1)}$ Q is the quality factor $Q = \frac{R_2\sqrt{1 + \frac{R_1}{R_2}}}{3R_2 + R_1 - KR_1}$ $\omega_o = \frac{\sqrt{1 + \frac{R_1}{R_2}}}{R_1 C}$
20-16G		$\frac{V_o}{V_i} = -\left[\frac{S^2 + \left(\frac{R_v - R}{R_v + R}\right)\left(\frac{\omega_o}{Q}S\right) + \omega_o^2}{S^2 + \left(\frac{\omega_o}{Q}\right)S + \omega_o^2}\right]$

$$C_1 = \frac{4}{3}C_2 \tag{20-25}$$

Substitution into Eq. 20-23 while invoking the necessary value of ω_0 leads to

$$2000 \text{ rad/s} = \frac{1}{\sqrt{3R_1^{\;2}C_2^{\;2} \times \frac{4}{3}}}$$

$$= \frac{1}{2R_1C_2}$$

or

$$R_1C_2 = 250 \times 10^{-6} \text{ s} \tag{20-26}$$

If C_2 is taken to be 0.05 µF, then Eq. 20-26 requires R_1 and hence R_2 to be 5 kΩ. Eq. 20-25 would then require

$$C_1 = \frac{4}{3}C_2$$

$$= \frac{4}{3} \times (5 \times 10^{-8} \text{ F})$$

$$= 0.066 \mu\text{F}$$

which is quite close to a readily available value of 0.068 µF. These are all reasonable values; hence, the example concludes with the circuit of Fig. 20-18.

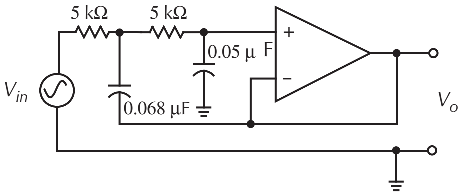


Figure 20-18. Second-order unity gain Bessel low-pass with a zero frequency group delay of 500 µs.

In fairness it is necessary to state that the solutions given to the two examples are not unique. In fact, even more elegant solutions than the ones given are possible though not as straightforward. These more elegant solutions will no doubt occur to the reader after further study.

20.3 Power Amplifiers

Power amplifiers for professional applications, unlike those intended for home entertainment use, must usually be capable of providing a multiplicity of voltage values at their output terminals. Furthermore, for reasons of safety and in order to avoid inadvertent mishaps in wiring or handling, it is often required that neither side of the output distribution lines be referenced to ground except in a balanced way through a high impedance in order to provide a static discharge path. These requirements are usually met by feeding the distribution lines from an isolated transformer secondary even though the transformer itself presents a source of distortion and bandwidth limitation.

Power amplifiers, when operated within their inherent limitations, are essentially constant-voltage sources. The sinusoidal rms voltage at the output terminals at rated power that is required in professional applications is commonly in voice coil values of 25 V, 70.7 V, or, in recent times, 200 V. The loudspeakers or other loads are, in the case of 25, 70.7, or 200 V lines, fed from the secondary of a stepdown transformer which has several primary taps for determining the actual average sinusoidal power supplied to an individual device. When feeding several devices from a common constant voltage distribution line it is only necessary to insure that the sum of the power taps to the individual devices does not exceed the output capability of the driving amplifier. High values, such as 70.7 V or 200 V, for the constant-voltage distribution system will minimize the I^2R loss in the distribution lines themselves (see Chapter 14). It is an absolute necessity, however, that the transformer at the amplifier, when such is employed, and the step-down transformers at the individual load devices be of high quality. Poor quality transformers with either high insertion losses and/or poor impedance characteristics will completely defeat the advantages offered by the constant-voltage distribution technique.

Many of the present-day power amplifiers intended for professional use are produced in two-channel versions even when the ultimate employment is to be with monaural program material. When preceded by active or passive crossover networks, such amplifiers can provide biamplification by devoting each amplifier channel to a separate part of the audio spectrum. This technique when properly employed may offer level adjustment, distortion, and loudspeaker damping advantages over the full-spectrum approach employing an individual amplifying channel. Additionally, such amplifiers may be employed with a balanced bridge output driven by both channels, which doubles the output voltage swing but requires a load impedance that is twice as large as a single channel alone. This technique can drive a 70.7 or even higher voltage balanced distribution line without a transformer at the amplifier but the ground isolation previously mentioned is lost in the process. This may furnish the user with a difficult choice.

Audio power amplifiers are designed in reverse, which means that the output stage is designed first followed by the design of the output driver stage that, in turn, is followed by the required intermediate stage or stages and then finally the input stage. Depending on the power, distortion, and efficiency requirements the class of operation of the output device or devices has traditionally been restricted to A, AB, B, or D. The most recent developments have widened the choice somewhat in that some current designs involve changing the supply voltage to the output stage under dynamic conditions. It appears that one may look forward to an entire alphabet of classes of operation. When a single device is employed in the output stage, class A operation, in which current exists in the active device throughout a complete cycle of signal swing, is the only acceptable class of operation. Class A is inherently the most linear class of operation. If pairs of output devices are employed in push-pull in the output stage, then classes A, B, AB, and AB plus B (at least two pair of devices) are distinct possibilities. Other than A, the other classes are in general more efficient but inherently are not as linear as class A. In class B operation each member of a push-pull pair is active over only one-half of a complete sinusoidal signal cycle. Class AB is intermediate in this regard between A and B. In AB plus B a pair of devices operates push-pull in class AB while a second pair of devices in push-pull operates nearly in class B. Class D is the designation given to the mode of operation wherein the output devices are operated in a switching mode. This means that the output devices are conducting as heavily as possible or not at all. This mode of operation offers efficiencies bordering on 90% but introduces a host of other problems with regard to radio-frequency interference as well as requiring specialized active devices, drive circuitry, and design techniques.

The advent of bipolar complementary symmetry transistors introduced the possibilities of many new circuit topologies in power amplifier design and the development of complementary symmetry power field effect transistors has opened up even more exciting avenues for truly superb amplifier developments.

Fig. 20-19 is a rather basic complementary symmetry bipolar transistor output stage for operation in classes A, AB, and B. The class of operation is dictated by the details of the biasing and drive arrangements.

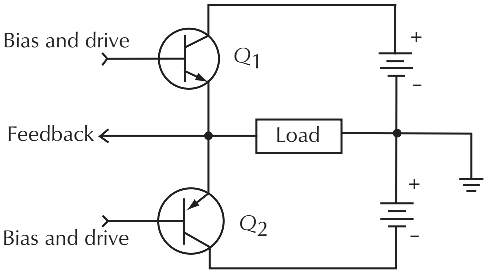


Figure 20-19. Complementary symmetry output stage.

The currents in Q_1 and Q_2 are equal at quiescence and there is no current in the load. If the bases of Q_1 and Q_2 are driven with a positive-going signal, Q_1 conducts more heavily while the current in Q_2 decreases, thus producing a net current in the load directed from left to right such that the left end of the load assumes a positive voltage relative to ground. On the other hand, if the bases of Q_1 and Q_2 are driven with a negative-going signal, Q_1 conducts less heavily while Q_2 conducts more, thus producing a net current in the load directed from right to left such that the left end of the load assumes a negative voltage relative to ground. The load in effect is connected in the emitter circuits of both transistors that consequently operate as common collector transistors. As is well known, the voltage gain of a common collector amplifier is slightly less than one and without polarity inversion. Hence, the driving circuitry must be able to produce a signal swing in excess of the swing to be expected across the load.

Fig. 20-20 represents a circuit configuration that is superficially similar to Fig. 20-19 but is drastically different in its operation. The currents in Q_1 and Q_2 are again equal at quiescence and there is no current in the load. When the bases of Q_1 and Q_2 are driven with a positive-going signal Q_1 conducts more heavily while the current in Q_2 decreases, thus producing a net current in the load directed from left to right such that the right end of the load assumes a negative voltage relative to ground. On the other hand, if the bases of Q_1 and Q_2 are driven with a negative-going signal, Q_1 conducts less heavily while Q_2 conducts more, thus producing a net current in the load directed from right to left such that the right end of the load assumes a positive voltage relative to ground. The load now in contrast to Fig. 20-19 has been shifted from the transistor emitters to the transistor collectors.

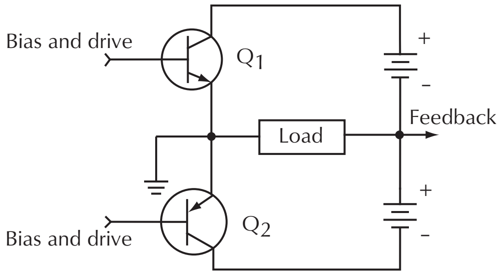


Figure 20-20. Common-emitter complementary symmetry output.

Instead of dealing with a common collector stage as in Fig. 20-19 the circuit of Fig. 20-20 is that of a common emitter amplifier. That is, the load is really in the collector circuits of the transistors. Such a stage produces an output signal swing that is inverted in polarity and possibly of much larger amplitude than the input signal swing. The drive requirements of such a stage are greatly relaxed as compared with those of the circuit of Fig. 20-19. In fact, if power field-effect transistors (FET) are employed instead of bipolar transistors in the circuit of Fig. 20-20, it is possible to produce a high-power high-voltage output stage which can be easily driven by a single low-power operational amplifier. This is made possible because a field-effect transistor is a voltage-controlled device having a high input impedance whereas a bipolar transistor is a current-controlled device with inherently a low input impedance in the common emitter configuration.

Fig. 20-21 is a totem pole configuration that has been supplemented by an additional pair of transistors to produce the AB plus B mode of operation. The totem pole configuration became popular before the advent of high-power complementary symmetry pairs and is still employed where it is desirable to use only a single type of power transistor. This configuration must be driven by a circuit that furnishes two drive signals of opposite polarity. Resistors R_1 and R_2 are typically a few hundred ohms while the resistors R_3 and R_4 are of the order of 1 Ω. At quiescence neither Q_3 or Q_4 is conducting and the emitter currents of Q_1 and Q_2 are equal in the range of 50–100 mA and there is no current in the load. If the base of Q_1 is driven with a positive-going signal while that of Q_2 is driven with a negative-going signal, the emitter current of Q_1 will increase while that of Q_2 will decrease and there will exist a net current in the load directed from left to right such that the left of the load will assume a positive voltage relative to ground. On the other hand, if the base of Q_1 is driven with a negative-going signal while that of Q_2 is driven with a positive-going signal, the emitter current of Q_1 will decrease while that of Q_2 will increase, thus producing a net current in the load directed from right to left such that the left end of the load will assume a negative voltage relative to ground. In the first instance, the current supplied to the load by Q_1 was forced to pass through R_3, and in the second instance, the current supplied by Q_2 is forced to pass through R_4. Under small or moderate signal swings the voltage drops across R_3 or R_4 are not sufficiently large to forward bias either Q_3 or Q_4. When larger voltage swings are occurring, Q_3 and Q_4 will be brought into conduction on alternate halves of the cycle and thus will aid in supplying load current. The circuit does possess a basic asymmetry in that even though Q_1 (and Q_3 when it conducts) are operated as common collector amplifiers, transistors Q_2 (and Q_4 when it conducts) are operated as common emitter amplifiers.

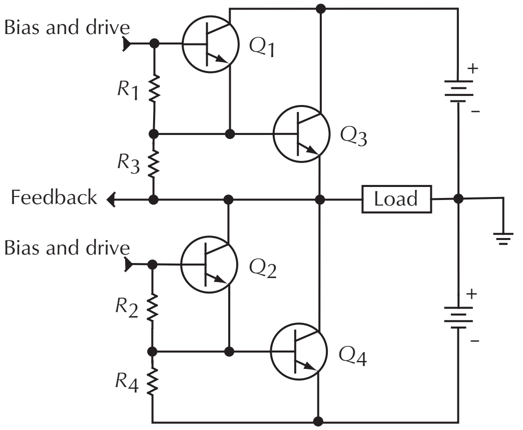


Figure 20-21. Totem pole with AB plus B.

Fig. 20-22 is a complete though modest amplifier. Q_1 and Q_2 are complementary symmetry monolithic darlington power transistors. Q_3 is employed to adjust the forward bias on the output stage. Q_4 is a constant current source that insures that Q_2 receives adequate voltage drive under large signal conditions. Most of the open loop voltage gain is provided by Q_5 while C_1 determines the dominant pole, which was discussed in connection with operational amplifiers. R_{12} and Z_1 determine the total current in the matched differential pair Q_6. The ac voltage gain of the amplifier is set by R_9 and R_{10} which determine the feedback fraction above a frequency of a few hertz. There is complete negative feedback at dc as brought about by the presence of C_2 in series with R_{10}. This in collaboration with Q_6 insures that the output voltage of the amplifier is zero when no signal is applied at the input. By monitoring the current in R_1 and R_2 it is possible to provide protection against load short circuits by means of a relatively simple additional circuit. If C_2, R_{10}, and R_9 were removed from the circuit of Fig. 20-22 altogether and further if the right- hand base of Q_6 were connected through a resistor equal to R_{14}, the resulting circuit would be a power operational amplifier built from discrete components. The present input would become the noninverting input while the right-hand base of Q_6 would be the inverting input.

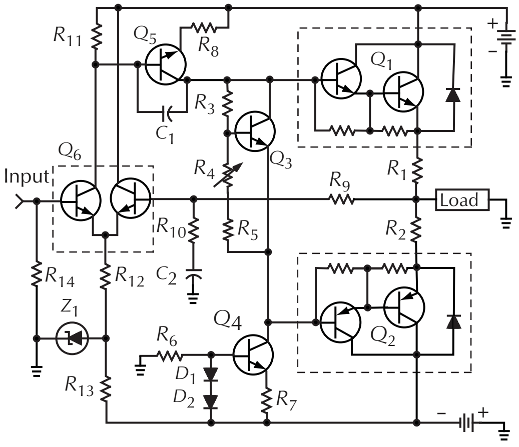


Figure 20-22. Complete power amplifier.

20.3.1 Protection Mechanisms

Power amplifier protection mechanisms fall roughly into two categories, the protection of the amplifier against faults in the load and protection of the load against faults in the amplifier. The amplifier designer must, unfortunately, shoulder the burdens of both categories. The load must be protected against turn-on and turn-off transients, against dc appearing at the amplifier output terminals unless that is the intended purpose of the amplifier, and against unwarranted oscillation in the amplifier caused by the load if the load itself presents a reasonable impedance to the amplifier. The amplifier must be protected against short-circuited or very low-impedance loads, excessive temperature within the amplifier, wide variations in ambient temperature, radio-frequency signals induced in the loudspeaker lines, radio-frequency signals induced in the input signal lines, dc on the input signal lines if such is not the intended use, and other types of reasonable abuse. All of the items above can be dealt with in practice but to do so involves an enormous additional expense in design, manufacture, and maintenance. Inferior as well as less costly products treat these features minimally if at all.

The most common of all load protection schemes is a fuse in series with the load. It may be a single fuse, fusing the overall system, or in a case of a multiway loudspeaker system, it may be one fuse on each loudspeaker.

Fuses help to prevent damage due to prolonged overload but provide essentially no protection against damage that may be done by large transients and such. To minimize this problem, high-speed instrument fuses such as the Littelfuse 361000 series should be used. Fig. 20-23 shows the fuse size versus loudspeaker power and impedance ratings.

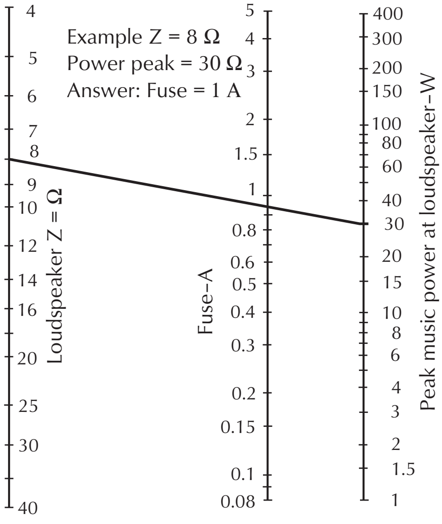


Figure 20-23. Fuse selector nomograph for loudspeaker protection. Courtesy Crown.

The load protection mechanism against turn-on, turn-off transients and against dc usually involves a pair of hard relay contacts energized by suitable circuitry. Upon turn-on these contacts are open and are subsequently closed by means of a delay circuit that allows the amplifier to stabilize before the load is connected. These same contacts open immediately upon amplifier turn-off. An additional signal is supplied to this muting circuitry by means of a low-pass filter connected to the amplifier's final stage. If dc is sensed at this point in excess of a safe value, the circuit acts to disconnect the load from the amplifier.

The amplifier can be protected against dc at its input when such is not intended by either transformer or capacitor high-pass filters. It may further be protected against radio-frequency signals at the input or output lines by means of series-connected low-pass filters. These filters must be designed with care so as to not unnecessarily restrict the intended amplifier passband. Excessive heat sink temperature is sensed by an attached thermal sensor that controls internal cooling fans or may ultimately interrupt power to the output stage. Thermally sensitive bias tracking circuitry can be provided to insure appropriate bias conditions for the output over a reasonable range of ambient temperature. Short-circuit protection usually involves monitoring the currents in the output devices and restricting the drive applied to the output stage whenever excessive current is detected with long-term protection still being provided by the thermal mechanisms previously mentioned. Such a circuit suitable for the amplifier of Fig. 20-22 is given in Fig. 20-24. Resistors R_{15} and R_{16} form a voltage divider sensing the emitter currents of the output devices. In the event of excessive emitter current, Q_7 robs base drive from Q_1 while Q_8 robs base drive from Q_2. Diodes D_3 and D_4 prevent Q_7 and Q_8 from having their collector to base junctions forward biased under conditions of normal operation. This same circuit can readily be converted into a dissipation limiter rather than just a current limiter by referencing the junction of the R_{16} resistors to ground rather than to the amplifier output terminal.

The amplifier depicted in Fig. 20-25 has an interesting protection mechanism that provides automatic turn-on muting along with protection of the output stage.

The output stage of this amplifier consists of complimentary MOSFETs connected in the common source configuration that yields voltage gain and signal polarity inversion in the output stage. When the amplifier is first energized any incoming signal is muted by the JFET T_1, which provides a low resistance to ground. As the capacitor connected to the gate of T_1 gradually charges toward −15 V this condition is relaxed and the amplifier becomes operative. During normal operation the currents in the output transistors T_2 and T_3 are monitored at points A and B. An excessive current in either device will trigger Q_2 which is a voltage discriminator or Schmitt trigger circuit. The amplifier will then be muted for a time determined by the *RC* combination in the gate circuit of T_1.

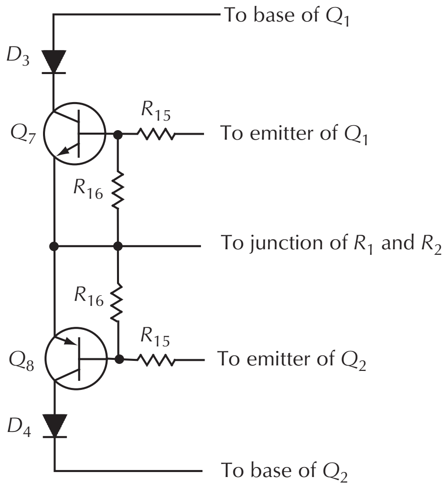


Figure 20-24. Short-circuit protector for the amplifier of Fig. 20-22.

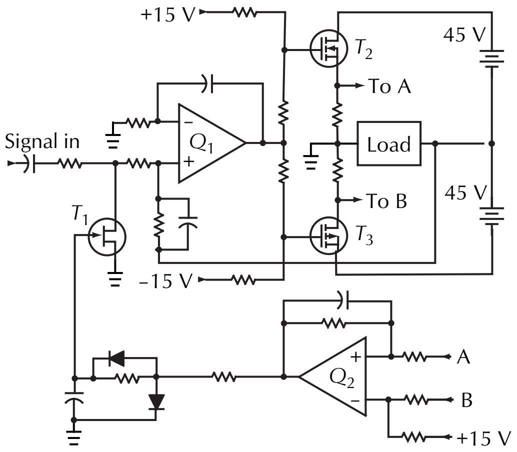


Figure 20-25. Protected MOSFET power amplifier.

20.3.2 High-Power Amplifiers

The demands of the sound reproduction and reinforcement industry for both higher power and higher performance amplifiers have brought about the necessity for a paradigm shift in power amplifier design. The older time-honored linear designs employing push-pull output stages in classes A, AB, or AB + B have poor output power efficiencies and/or appreciable internal power dissipation at quiescence. Class A dissipates its rated power internally at quiescence and only approaches a power efficiency of 50% when delivering its full output power. A pure class B stage would have zero internal quiescent power dissipation and a power efficiency approaching 78.5% only at full output while being plagued with unacceptable crossover distortion. Class AB solves the crossover distortion problem while introducing some internal quiescent power dissipation along with a smaller output power efficiency than pure class B. Class AB + B retains the low quiescent power dissipation of class AB and approaches the power efficiency of pure class B when operated at full output. Such amplifiers most often employ conventional power supplies consisting of a power transformer whose primary is energized from the ac mains and whose secondary is applied to a full wave bridge rectifier connected to a capacitor input filter. This arrangement can also yield bipolar dc supplies when the secondary of the transformer is center tapped and two capacitor filters are employed. Large values of capacitance must be employed, as the fundamental ripple frequency is 120 Hz. Such supplies inherently suffer from poor voltage regulation and demand excessive root mean square (rms) current draw from the ac mains as their power factors fall in the range of about 0.6 to 0.7. These power amplifiers most often employ complementary symmetry bipolar junction transistors and ordinarily have power limitations that are dictated by the voltage breakdown properties of the active devices when conventionally employed. Some clever schemes for surmounting this limitation will be discussed subsequently. The linear power amplifier designs discussed above might well be referred to as being analog high-power amplifiers.

The paradigm shift necessary to achieve even higher output power and performance has affected the design philosophy of both the amplifier power supply as well as that of the power amplifier itself. Rather than employing continuous or analog techniques, the really high-power units now employ switching techniques in both the power supply and in the amplifier circuitry. The concept of having a pair of output devices, one positive and one negative, each alternately being either full on or full off, is not new. Audio power amplifiers employing solid-state active elements operated under what is termed *class D* have been around since about 1970. The less than spectacular performance of the early efforts was not a result of the failure in operating principle but rather the result of the shortcomings of the available active elements involved. These shortcomings have been diminished in modern power MOSFETs and IGBTs to the point that switching amplifiers are not only viable but also desirable in high-power applications. Additionally, switching topologies exceeding the properties of the classic class D have also been evolved. These will be discussed in a subsequent section.

Class D switching amplifiers have the desirable property that the output efficiency can approach 100% independent of the operating power level. This results because the active output devices ideally are either fully conducting or fully not conducting—i.e., they act as switches. This is best explained by reference to Fig. 20-26, which is a functional diagram of a class D amplifier, and Fig. 20-27, which displays pulse width modulation waveforms.

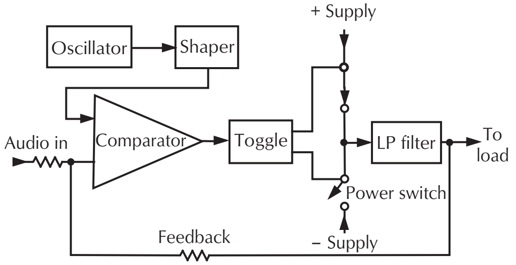


Figure 20-26. Functional diagram of a class D amplifier.

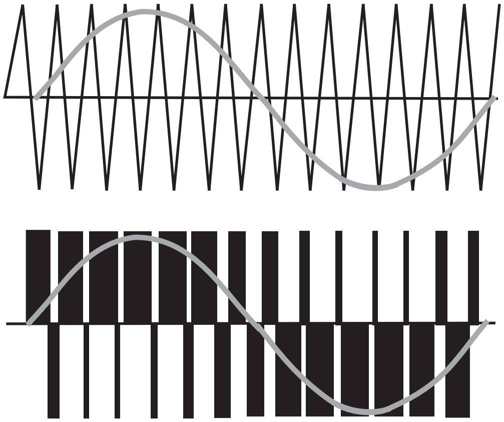


Figure 20-27. Pulse width modulation waveforms.

In Fig. 20-26, the continuously running oscillator operates at a fixed amplitude and with a fixed frequency usually between 200 kHz and 500 kHz. The shaper converts this signal into a triangular waveform at the same fundamental frequency and supplies the resulting waveform as one input to the comparator. The other input to the comparator is the summing node of the audio input signal and the feedback signal, which in the language of operational amplifiers means it is the error signal. For proper operation, the peak-to-peak value of the error signal must always be less than the peak-to-peak value of the triangular waveform at the other comparator input. This combination of the oscillator, shaper, and comparator forms a pulse-width modulator. Although the output of the comparator swings alternately positive and negative, the time spent in these excursions is in general different. The output pulse width of the modulator is proportional to the error signal. This results from the fact that the peak-to-peak value of the error signal is less than the peak-to-peak value of the triangular waveform and that the polarity from the comparator reverses depending on whether the instantaneous value of the error signal is greater or less than the instantaneous value of the triangular waveform. In the variable duty cycle, positive and negative pulses from the modulator toggle the active output devices depicted as switches either full-on or full-off. Thus, constant positive or negative voltage pulses are applied to the low-pass filter and load for variable intervals of time. The low-pass filter passes the time average value of these pulses in the audio band, producing a voltage across the load proportional to the instantaneous value of the original incoming audio signal. The high efficiency stems from the fact that there is no power dissipated in an output device when it is nonconducting and very little power dissipated when it is conducting in saturation, or fully on, as even though the current through the device is large, the voltage drop across the device is very small. Unfortunately, the output devices are highly specialized in that they must exhibit very fast switching times with the absence of charge storage effects. This alone pretty well rules out bipolar power transistors capable of handling the large currents and sustaining the high voltages involved. Vertically structured MOSFETs are usually employed as the switching elements in this basic simple design. Two other drawbacks to this simple class D structure are the possibility of radio-frequency generation and the difficulty of optimizing the low-pass filter for use with more than one value of load resistance. Loudspeakers are hardly constant impedance devices, much less constant resistance devices. The most recent designs in switching amplifiers have addressed these problems.

20.3.3 High-Power Analog Amplifiers

Crown International, principally through the work of Gerald Stanley, is responsible for a continuing series of technical innovations in high-power analog amplifiers involving class AB + B that was first introduced by Crown. Though not changing the basic efficiency of this configuration, the innovations have led to high-powered designs up to several thousands of watts in individual units. The innovations involve output stage topology, amplifier cooling, and power transistor safe operating area assessment as well as control. At the heart of these innovations is an output stage topology that is a full bridge configuration with one output terminal always at ground potential. A simplified view of this configuration is given in Fig. 20-28.

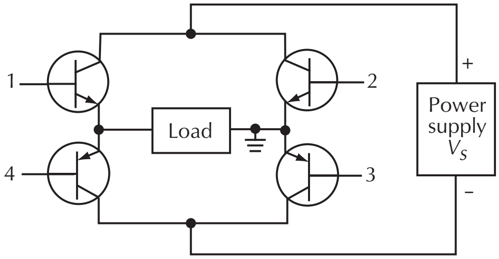


Figure 20-28. Crown-grounded full bridge topology.

In Fig. 20-28 the transistors at points 1, 2, 3, and 4 represent composites of several NPN and PNP bipolar power transistors constituting AB + B arms of the bridge. The NPN transistors at 1 are the positive voltage output stage while the PNP transistors at 4 are the complementary negative output voltage stage. The NPN transistors at 2 are the positive bridge balance output stage and the PNP transistors at 3 are the complementary negative bridge balance output stage. When a positive output is required, the transistors at 1 conduct connecting the left end of the load to the positive terminal of the supply and the transistors at 3 conduct connecting the negative terminal of the supply to ground. When a negative output is required, the transistors at 2 conduct connecting the positive terminal of the supply to ground and the transistors at 4 conduct connecting the left end of the load to the negative terminal of the supply. The control, bias, and driving circuitry must ensure that at quiescence there is no voltage drop across the load and, when delivering a signal, that the voltage division is correct across the diagonally opposite pairs of transistors that are driven toward non conduction. This arrangement offers two very distinct advantages as compared with a conventional complementary symmetry output stage: it requires only a single power supply voltage V_S to produce a peak-to-peak voltage swing across the load equal to $2V_S$ and it simultaneously halves the sustaining voltage requirements of the output devices.

The composite transistors at 1 and 2 are mounted directly without electrical insulators in order to ensure good thermal contact to one heat sink that itself is electrically insulated. The composite transistors 3 and 4 are mounted directly to a second electrically insulated heat sink. The heat sinks themselves, rather than the usual heavy metal extrusions, involve many thin metal fins such as those employed in refrigeration and air conditioning technology or automobile radiators. This greatly increases the surface area exposed to a forced air stream for cooling purposes and allows the power devices to operate with lower junction temperatures than would otherwise be the case.

The final contributor to high-power operation of this circuitry involves what amounts to a small, dedicated analog computer that emulates the operation of the output stage consistent with the operating conditions that exist in the amplifier in real time. This allows the output stage control circuitry to restrict the drive to the output such that the output devices remain always in what is the safe operating area of the moment.

Another approach to high-power operation, which also offers an efficiency advantage over class AB while basically employing AB circuitry, involves changing the supply voltage from a nominal value to a higher value when larger output swings are called for. This can be accomplished with a fixed output stage configuration powered by a variable switching power supply, by switching the supply voltage from a fixed nominal value and a fixed higher value, or by using two sets of output devices with one set powered by a nominal voltage supply and the other set powered by a higher voltage supply and switching between the sets of output devices. This last method of operation is termed *class G*. QSC Audio was the first manufacturer to introduce such an amplifier for employment in the sound reinforcement industry. This design can still fall in the analog category as the switching between the sets of output devices is purely accomplished by analog techniques as devised by Pat Quilter of QSC. Fig. 20-29 is a simplified schematic of the positive half of a complementary-symmetry class G output stage that illustrates the principal of operation.

For small-amplitude drive signals, the emitter follower in the lower part of Fig. 20-29 is powered by the supply with voltage V and operates as a normal AB stage consistent with this value of the supply voltage. The upper transistor is not forward biased and is not conducting so the supply with $2V$ plays no role. For large amplitudes of the drive signal, when the value of the drive signal approaches V, the lower transistor approaches saturation and sufficient forward bias is applied to the upper transistor to bring it into conduction, thus bringing into play the supply of voltage $2V$. For even greater values of drive voltage, the supply of voltage V is disconnected by the diode on the right because this diode is now reverse biased and the lower transistor is in full saturation. Under this condition, the upper transistor is operating as an emitter follower with a supply of voltage $2V$. The operation of the negative half of the output stage is the same except all polarities are reversed.

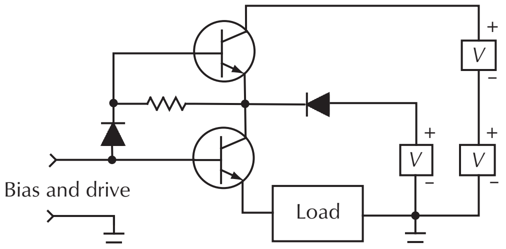


Figure 20-29. Simplified positive half of a class G output stage.

A further step in this same direction also employed by QSC as well as other manufacturers is that of the class H topology for the power stage. Instead of two sets of devices permanently connected to two different voltage supplies on the positive half as well as the complementary negative half of the output stage, class H employs a single set of devices on the positive half and a complementary set on the negative half. The supply voltages to these devices are switched to different values according to the requirements of the audio signal at the moment. Such an arrangement is illustrated in Fig. 20-30.

In Fig. 20-30 the positive and negative class H output stages usually consist of paralleled bipolar power transistors and an associated driver with the negative half being complementary with the positive half. The class of operation of each half is basically that of class B except for the very lowest signal levels. Efficiency is improved by maintaining as low as possible voltage drop across the active devices when they are delivering current to the load. Consider for the moment that the output signal is swinging positive and its instantaneous value is approaching the fixed value of $+V$. Switch $S+$ is a voltage comparator–operated switch. This switch is closed when the output signal exceeds a fixed positive reference voltage whose value is chosen so that the stage never goes into saturation. With $S+$ closed, the output signal can now increase if required up to slightly less than the rail limit of $+2V$ whereas the voltage across the active transistors themselves is always less than $+V$.

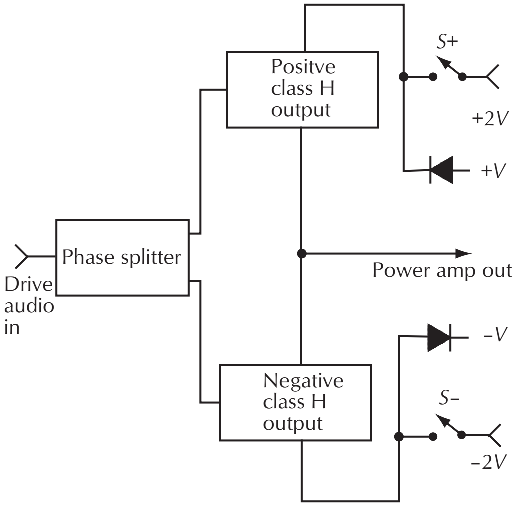


Figure 20-30. Basic class H output stage.

The operation of the negative half mirrors that of the positive half for negative swings of the output signal. One is not restricted to the employment of just four voltage supply values consisting of two positive and two negative. One might in fact employ three or more on each side by increasing the number of fixed voltage taps on the power supply, adding more comparator-operated switches, and reverse-biased diodes as appropriate for each switch. In this fashion, the voltage supply of the moment more accurately tracks the instantaneous requirements of the signal and the dissipation in the active output devices is kept near a minimum.

20.3.4 Switch Mode Power Supplies

Switching power supplies are now employed in practically all high-power audio amplifiers be they of the analog type of classes AB + B, G, and H or of the various manifestations of switching amplifier stemming from class D. A variety of considerations have compelled this change from former practice. When the power requirements are large, a switching power supply offers significant size and weight advantage over a conventional supply employing an ac-main-operated power transformer, full-wave bridge rectifier, and capacitor filter bank. More importantly, a switching power supply offers the advantage of active power factor correction. This latter feature is crucial for obtaining the maximum power from a single 120 Vac outlet having limited current capability. Additionally, most switching supplies feature a convenient selection feature that allows operation from either 120 V or 240 V mains. In fact, the more advanced designs of switching supplies can operate over an ac voltage range of 90 V to 270 V rms without any internal circuit changes.

These desirable features of switching power supplies come with a price in that the design of such units is highly specialized and much more engineering effort must be expended in yielding a viable unit. High-power units require auxiliary supplies for the control circuitry that must be fully active during main supply startup as well as normal operation. The fundamental switching frequencies fall in the range of 30 kHz to 100 kHz with the switching pulses being pulse width modulated. As a result, harmonics are generated in the radio-frequency range. This property necessitates the incorporation of a sophisticated electromagnetic interference (EMI) filter to prevent the appearance of the switching frequency and its harmonics as common-mode signals on the ac supply lines. Fig. 20-31 is a block diagram of a typical switch-mode power supply for use with high-power audio amplifiers.

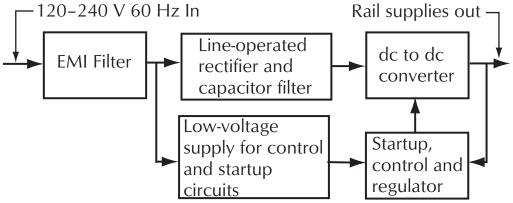

120–240 V 60 Hz In
Rail supplies out
EMI Filter
Line-operated rectifier and capacitor filter
dc to dc converter
Low-voltage supply for control and startup circuits
Startup, control and regulator

Figure 20-31. Block diagram of switch mode power supply.

The EMI filter must be designed to prevent common-mode signals that are generated in the power supply from being conducted to the external supply mains while at the same time offering minimum series impedance to the differential 60 Hz ac voltage of the supply mains. This is accomplished by having balanced inductors in each of the supply lines with positive mutual inductance between the inductors in upper and lower lines. This arrangement maximizes the inductance for common-mode currents that flow in the same direction in both conductors. For the oppositely directed differential currents of the 60 Hz main supply, the mutual inductance is negative, thus forcing the overall series inductance to a small value. A typical EMI filter circuit is shown in Fig. 20-32.

The line operated rectifier and capacitor filter immediately follows the EMI filter. Such an arrangement appears in Fig. 20-33.

The circuit of Fig. 20-33 features an inrush current limiter in the form of a thermistor that presents a large

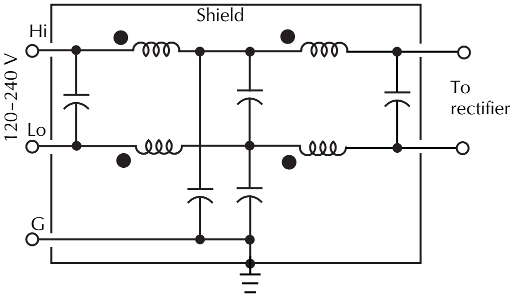


Figure 20-32. Typical EMI filter.

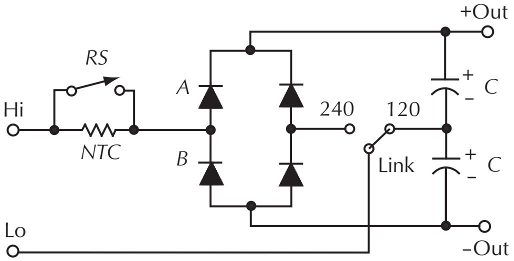


Figure 20-33. Line-operated rectifier and capacitor filter.

resistance when cold and a decreasing resistance as the device heats up from the passage of current. This behavior limits the peak current drawn from the line when the circuit is first energized from a cold start. After the capacitor bank is fully charged the thermistor is shorted out by a relay-controlled switch as part of the start-up protocol of the dc-to-dc converter. With the link connected as shown, the storage capacitors act as voltage doublers. When the Hi side of the line is positive, diode *A* charges the upper capacitor with the indicated polarity. When the Hi side of the line goes negative one-half cycle later, diode *B* charges the lower capacitor with the indicated polarity. Voltage doubling occurs because the two capacitors are permanently connected in series. For a 240 V supply line, the link is moved to the 240 V position and the circuit is then that of a normal full-wave bridge rectifier having an effective capacitance of 0.5*C*. The nominal total dc output voltage is the same in either case.

The core of any switch-mode power supply is the dc-to-dc converter. Switch-mode supplies that are called on to deliver significant power usually employ either a half-bridge or full-bridge converter with the full-bridge converter being favored for employment in the supplies of the most powerful audio amplifiers that make use of full switching technology in their design. A greatly simplified diagram for such a full-bridge converter is exhibited as Fig. 20-34 from which the basic operation may readily be understood.

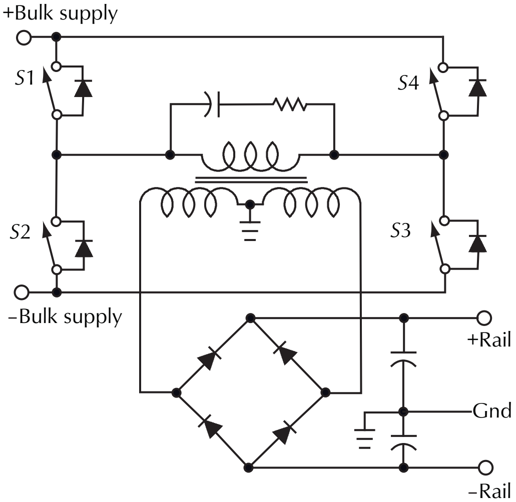


Figure 20-34. Full bridge dc-to-dc converter.

In Fig. 20-34 the switches *S*1 through *S*4 represent either insulated gate bipolar power transistors or N channel power MOSFETs. On the first half of the switching cycle, switches *S*1 and *S*3 are closed in concert and connect the primary of the transformer across the bulk dc supply so that current flows in the primary from left to right for some variable period of time. On the second half of the switching cycle with *S*1 and *S*3 now open, switches *S*2 and *S*4 are closed in concert and connect the primary of the transformer across the dc bulk supply so that current flows in the primary from right to left again for some variable period of time. The switches are activated by the control circuitry and by varying the duty cycle of the switches it is possible to maintain the level of the rail voltages in spite of varying rail loads and variations in bulk dc supply values. The secondary of the transformer is center tapped and feeds a full-wave bridge rectifier and a capacitor filter for each rail. The capacitance values required here are modest as compared with those in the bulk supply as the ripple frequency is twice the switching frequency and can range from about 60 kHZ to 200 kHz. The *RC* network in parallel with the transformer primary is a snubber network that in conjunction with the diodes in parallel with the switches allows switching transients to be damped while returning energy to the bulk supply.

20.3.5 Technological Innovation

The most recent high-power analog amplifier design is based on a patented design that introduces a new class of operation. This design is based on an innovation that truly represents the expression, “thinking outside of the box.”

The innovation put forth by the Swedish firm Lab.gruppen has resulted in the class TD, which is a tracking class D amplifier. The objective was to produce a very powerful audio power amplifier in which the entire signal path is that of an analog-type amplifier. This is accomplished by feeding the original audio signal to both an analog amplifier as well as a classic class D amplifier in parallel. The class D amplifier is structured in the form of a half-bridge as depicted in Fig. 20-26 except that instead of having a loudspeaker as the load on this amplifier, it is the output stage of the analog amplifier that serves as a load of the class D amplifier. The effect of this arrangement is to produce a class AB + B amplifier that has continuously varying positive and negative rail supplies that track the needs of the audio signal of the moment. The output stage of the analog amplifier operates with high efficiency as the voltage across the output devices at any instant exceeds the amplifier output voltage only by the amount necessary to insure active operation of the devices. This topology is illustrated in simplified form in Fig. 20-35. The overall system is powered by a switch-mode power supply. This design is said to retain the distortion and noise characteristics of an analog amplifier while closely approaching the output efficiency of a class D amplifier.

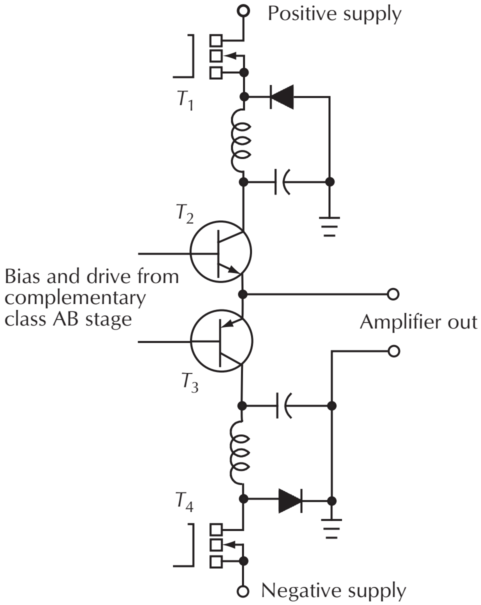


Figure 20-35. Class TD or tracking class D amplifier.

In Fig. 20-35 T_1 and T_4 represent several parallel N channel power MOSFETs that represent the switching elements of a class D amplifier each with an associated low-pass filter typical of class D operation. Sandwiched in between these two parts of the half-bridge are the output devices of the analog amplifier that constitute the load for the class D amplifier. T_2 and T_3 represent several parallel complementary NPN and PNP bipolar power transistors connected as emitter followers that directly drive the loudspeaker load.

20.3.6 Class D in Full Bloom

Instead of trying to skirt the problems associated with class D, Gerald Stanley and his design group at Crown Audio, Inc. have faced them head-on and over a period of time have evolved wide-ranging innovative solutions from the ac power cord to the amplifier output terminals.

The innovations begin in the switch-mode power supply that can operate, without any necessity for internal changes, from line sources ranging from 85 Vac to 277 Vac 50–60 Hz. Full power is obtained for supply voltages ranging from 120 Vac to 240 Vac. The bulk dc supply is a full-wave bridge rectifier and capacitor filter combination. This is followed by a unique dc-to-dc converter consisting of two half-bridges that are operated in a novel way in that the high-frequency switching signals to the two halves of the converter are phase shift modulated. This mode of operation impresses across the primary of the high-frequency step-up transformer an alternating square wave voltage whose duty cycle can be varied from zero to 50%. The rail voltages are derived through full wave rectification and filtering of the secondary high-frequency voltage from the step-up transformer. The positive and negative rail voltages can range from zero to some maximum value as the duty cycle of the primary square wave ranges from zero to 50%. A control loop constantly compares the rail voltage and current with reference values and adjusts the duty cycle in such a fashion that voltage regulation of the rail supply output is maintained for both changes in load as well as raw supply and all the while maintaining an overall power factor close to 0.95. Protection mechanisms are included to handle internal amplifier or overload problems.

The output stage is a far cry from the classic class D design as it employs an innovative topology as well as an advanced form of pulse width modulation. The output stage topology is termed *BCA*, standing for balanced current amplifier. Fig. 20-36 is a bare bones illustration of the BCA output stage.

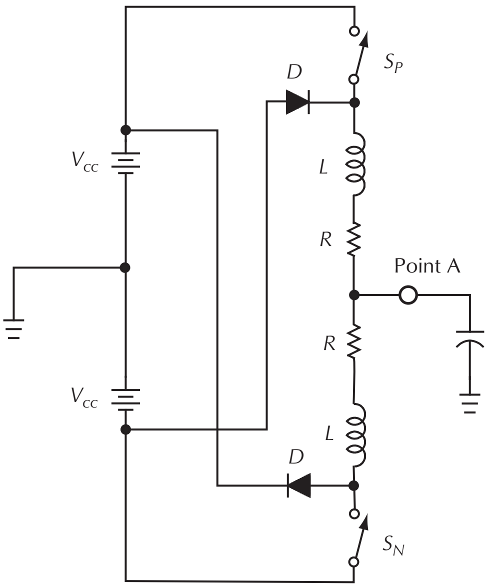


Figure 20-36. Simplified balanced current output stage.

In Fig. 20-36 the batteries represent the positive and negative rail supplies, S_p and S_n represent two groups of several paralleled N channel power MOSFETs that serve to connect to the positive and negative rail supplies, respectively, and there are two matched inductors as well as two matched diodes. An understanding of the operation of this output stage and its associated modulation technique may be had by examining three conditions or modes of employment: quiescent mode when no audio output is called for, positive mode when a positive-going audio output is required, and negative mode when a negative-going audio output is required. Recall that one component signal of a pulse width or switching modulator is a triangular waveform. The fundamental frequency of the triangular waveform employed in the associated modulator for this output stage is 250 kHz so the full period is 4 microseconds (4 μs) and the half period is 2 microseconds (2 μs). In the quiescent mode when no audio output is required the switches S_p and S_n are closed and opened in concert. Both switches are closed for the first half-period or 2 μs and both are open for the second half-period also of 2 μs. The inductors in Fig. 20-36 have relatively large self-inductance and quite small resistance such that their time constants, being the ratio of L/R, are very much larger than 2 μs so that the current in each inductor starts from zero and grows linearly in the clockwise sense around the closed loop at the rate of V_{CC}/L when the switches are simultaneously closed. There is no current in the diodes as they are reverse biased under this condition. All during the growth period, the voltage drop across the upper L/R series combination and the lower L/R series combination are each equal to V_{CC} while the voltage between point A and ground remains at zero. All during this time energy is being stored in the magnetic fields associated with each inductor and this stored energy reaches a maximum value when the elapsed time reaches 2 μs. When the elapsed time reaches 2 μs both switches are opened simultaneously and the circuit effectively becomes the depiction presented in Fig. 20-37.

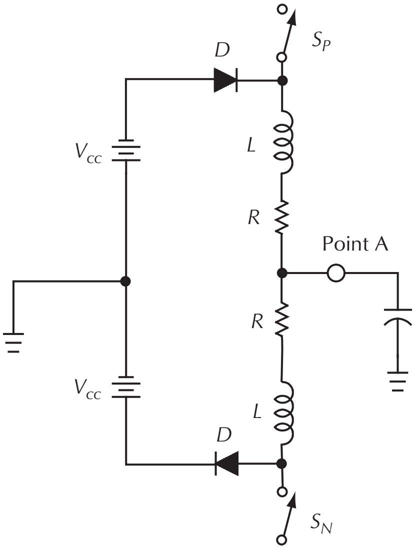


Figure 20-37. Equivalent decay circuit for current.

When the switches are simultaneously opened, the current begins to ramp down from the maximum value achieved in the previous 2 μs when the switches were closed and the diodes that played no role in the previous 2 μs maintain circuit closure but now to the opposite terminals of the batteries. In the next 2 μs the collapsing magnetic fields of the inductors continue to drive the current with a linearly decreasing magnitude and energy is being returned to the rail supplies. The energy recovery is not quite complete as there is a small heat loss in the imperfect inductors and diodes. During the ramp down process, the voltage from point A to ground again ideally remains at zero. Thus, in the quiescent state with no audio signal present, the voltage at point A remains at zero and there is no ripple arising from the modulation process.

When a positive-going audio signal is present the operation may be understood by reference to Fig. 20-38.

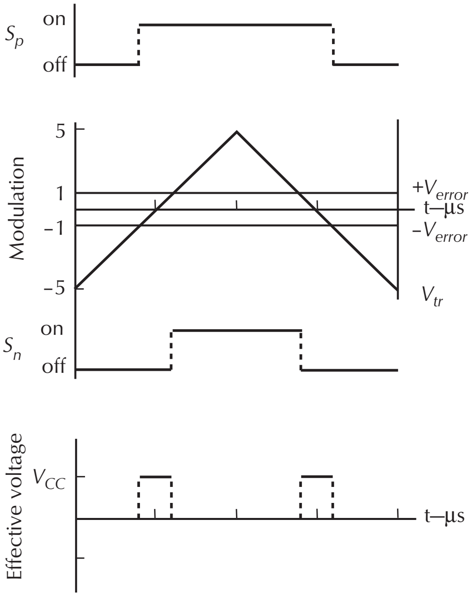


Figure 20-38. Modulation, switching, and effective voltage waveforms when the instantaneous audio signal calls for a positive 20% of full-scale output.

The three sketches in the top portion of the figure describe, respectively, the operation of the positive switch, the modulator behavior, and the operation of the negative switch while the final sketch at the bottom displays the effective voltage presented to the output filter of the BCA. Observe that the positive switch is turned on when the triangle waveform first crosses the negative error voltage value and stays on until the triangle waveform again crosses the negative error voltage value. Note also that the duration of the on-time of the positive switch is 2.4 µs. Observe also that the negative switch is turned on when the triangle waveform first crosses the positive error voltage value and stays on until the triangle waveform again crosses the positive error voltage value. The duration of the on-time of the negative switch is 1.6 µs. When both switches are off there is no output and when both switches are on simultaneously there is no output. There are two periods when the positive switch is on while the negative switch is off during which output is generated. Thus, the fundamental ripple frequency of the output is twice that of the fundamental frequency of the triangle waveform or 500 kHz. The switches themselves operate at only the fundamental frequency of the triangle waveform or 250 kHz. Another very important observation is that the on time of the two switches adds to 4 µs, which is the period of the triangle waveform. This is always true for this type of modulation technique. If an even larger output had been required, the negative error voltage line would have been lower while the positive error voltage line would have been higher. In such an instance, the on period of the positive switch would be increased while that of the negative switch would be correspondingly reduced. The two positive output pulses would still have amplitudes of V_{CC} but each would endure for a longer period such that the average value following the output filter would be correspondingly larger.

When a negativegoing audio signal is present, reference must be made to Fig. 20-39.

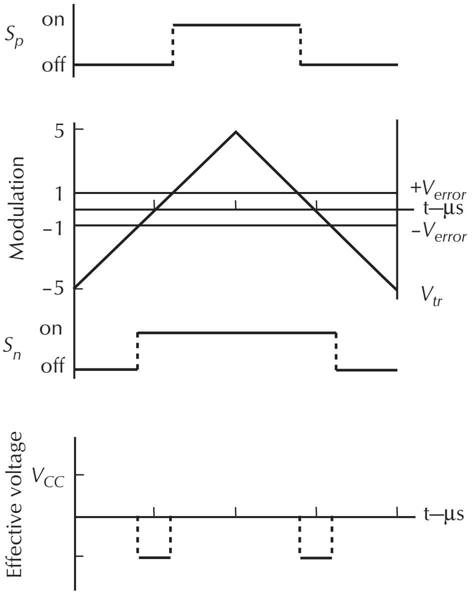


Figure 20-39. Modulation, switching, and output waveforms when the instantaneous audio signal calls for a negative 20% of full-scale output.

A study of Fig. 20-39 indicates that now the negative switch is turned on when the triangle waveform first crosses the negative error voltage value and remains on until the triangle waveform again crosses the negative error voltage line. In brief, the role of the two switches has now been reversed. The period of the negative switch is now 2.4 µs while the period of the positive switch is now1.6 µs. The output pulses as a consequence are now negative. The other general properties remain the same.

The output of the BCA under all three modes of operation—quiescence, positive audio output, and negative audio output—is further clarified by viewing the equivalent circuit displayed in Fig. 20-40.

When the three-position switch in Fig. 20-40 is in position two as shown in the drawing, the output of the

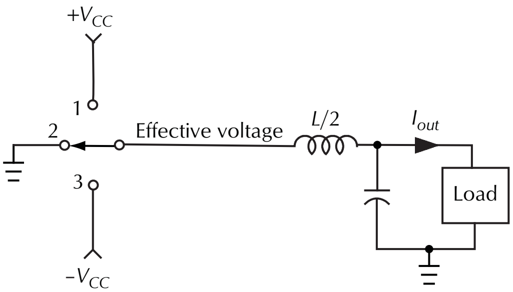


Figure 20-40. Equivalent circuit of BCA for all three modes of operation.

BCA is zero. This condition corresponds to the times when S_p and S_n are simultaneously both on or both off. Position one of the three-position switch corresponds to those intervals of time when S_p is on and S_n is simultaneously off. This position generates a positive audio output to the load. Finally, position three of the three-position switch corresponds to those intervals of time when S_n is on and S_p is simultaneously off. This final position occurs when a negative audio output is being generated.

The type of pulse width modulation employed with the BCA is called natural double-sided interleaved with $n = 2$. The n designator indicates the fundamental ripple frequency of the BCA output when generating audio signals relative to the fundamental frequency of the triangle waveform. The advantage of this type of modulation in addition to the fundamental ripple frequency being twice that of the triangle waveform is that there are no harmonically related distortion products to the audio frequency being processed and there are no odd integer multiple bands related to the fundamental frequency of the triangle waveform. As a consequence, only a relatively simple output filter is required for handling normal loudspeaker loads. Additionally, when operating a two-channel amplifier in the full-bridge mode, if the modulator in the second channel is in quadrature with that of the first, n becomes 4 rather than 2 so that the lowest ripple components appear at 1 MHz rather than at 500 kHz. Crown terms such amplifiers as being *opposed current interleaved amplifiers)*OCIA). This constitutes a new class of amplifier or class I. This type of operation is incorporated in Crown's I-Tech series of switching power amplifiers.

The required modulator circuitry is actually quite simple as is illustrated in Fig. 20-41.

In reference to Fig. 20-41, A_1 is the error signal amplifier while A_2 is a simple inverter amplifier. C_1 and C_2 are high-speed comparators.

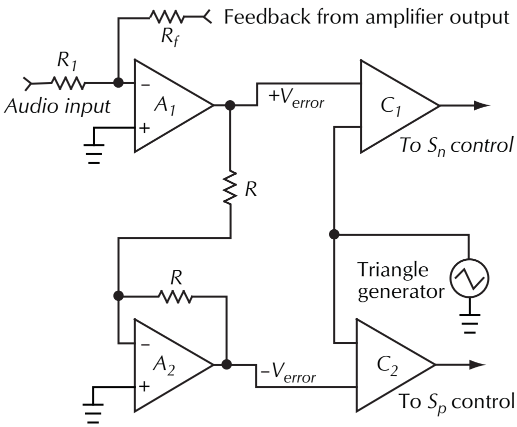


Figure 20-41. Basic OCIA pulse width modulator.

The innovations described in this section have led to the production of switching amplifiers having power ratings up to 8000 watts while producing only 0.35% of total harmonic distortion at rated power. The size and weight figures are equally as impressive being only 2 rack units and 29 pounds.

20.3.7 Signal Processing in Power Amplifiers

Most high-power audio amplifiers for professional applications currently feature a wide variety of signal processing functions. This was not always the case. Originally power amplifiers were just what the name implied with the possible exception of a selectable high-pass filter at the input for the protection of high-frequency compression drivers when such loudspeaker elements constituted the only load on the amplifier. During this period, full-range loudspeaker systems employed passive dividing networks and high-powered systems featured compressors contained in dedicated units preceding the power amplifier. The modest consoles of this period usually provided only high and low shelving filters. Modern changes began to occur first with the introduction of ⅓-octave real-time analyzers and dedicated ⅓-octave equalization units. These changes were further accelerated by the advent of TEF and similar computer-based analysis systems. Electronic crossover networks, signal alignment, and bi- or triamplification shortly came into vogue in order to correct system problems discovered by the new sophisticated analysis systems. Two-channel power amplifiers with the ability to operate the two channels independently or in the bridged mode became the de facto standard. The electronic crossovers and signal delays of this period were dedicated separate units and initially were based on analog techniques and slowly gave way to digital techniques as this newer technology developed. Digital techniques came into full flower with the development of powerful digital signal processing chips or DSPs. It now became economically feasible to concentrate the functions of filtering, equalization, signal delay, compression or limiting, and frequency division or crossover into a single system interposed between the input console and the power amplifier. This new digital signal processing unit was given different names by different manufacturers but was popularly referred to as a *loudspeaker management system*. The current trend in more powerful power amplifier design is to provide the entire required signal processing function within the confines of the amplifier chassis itself. In many instances there is no requirement for an external loudspeaker management system. These amplifiers usually accept audio in both analog and digital formats with the digital format conforming to the AES3 standard. If the amplifier is to be employed in live sound reinforcement, a premium is placed on having low latency in the amplifier's digital signal processing functions. Latencies of 2 µs or less are highly desirable.

20.3.8 Computer-Controlled Power Amplifiers and Systems

Computers are permeating every field of human endeavor and audio systems are no exception. Crown International Electronics in the middle 1980 s pioneered computer control to the depth of the level of individual amplifiers. Crown, building on the digital electronics and programming experience it had acquired in the development of the TEF analyzer, began designing a new line of power amplifiers that would be amenable to both digital control and digital monitoring with the incorporation of a plug-in digital module. Parallel to the development of these amplifiers, Crown also developed the computer interface and communication system necessary for interaction with these amplifiers. This work culminated in the Crown IQ system.

The original Crown IQ system was structured on three levels with microprocessors at each of the levels. At the uppermost level were the host computer and the IQ system software. The host computer could be any computer that had a serial (RS232, RS422, or RS4230) port. The computer with the installed IQ system software acted as a monitoring and control station. At the intermediate level was the Crown IQ interface that served as a communication device between the individual power amplifiers and the host computer. At the lowest level were the individual power amplifier plug-in microprocessor cards that were connected in a daisy chain by means of a single twisted pair to form a serial loop to the IQ interface as shown in Fig. 20-42.

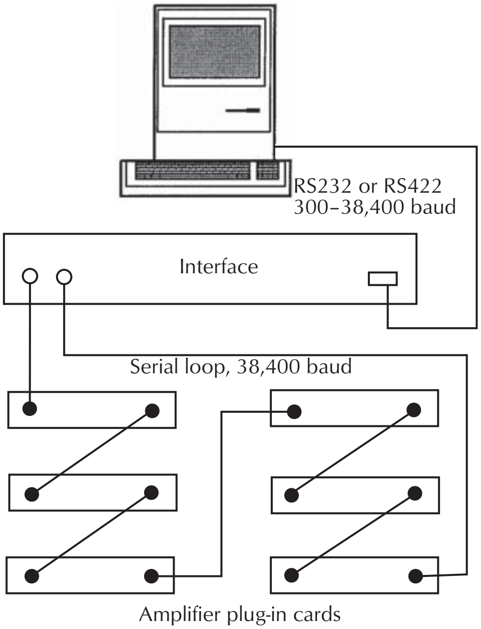


Figure 20-42. Basic OCIA pulse width modulator.

Communication between the interface and the individual amplifiers was at a baud rate of 38,400 so that the system operation appeared to occur almost in real time. All of the normally manually controlled functions of each amplifier could be computer controlled by this system up to a total of 2000 two-channel power amplifiers. The outstanding feature of this approach, however, was that the actual operational status of each amplifier including on-off, input level, output level, distortion, and safe operating area were constantly monitored almost in real time. Subsequently, the Crown system was expanded to include plug-in modules incorporating digital signal processing capability in addition to the original features. This allowed, in addition to the original monitoring and control features, online filtering, signal delay, and equalization, all under computer control.

The modern era of control of sound systems began in 1993 with the step taken by Peavey Electronics Corporation in the introduction of the MediaMatrix® System that featured real-time network communication and control of digital audio signals by Ethernet supported by a CobraNet® hardware interface. This system allowed the interconnection, communication between, and control of all of the elements associated with even the most sophisticated sound systems.

Crown's original IQ system has evolved to the current form termed HiQnet™ that features communication via serial, Ethernet, USB, and CobraNet® audio. This current system not only links, controls, and monitors power amplifiers but also the other elements associated with the smallest or largest of sound systems as well, all under computer control.

Practically all power amplifier manufacturers currently feature some form of computer networking and control of their products.

Bibliography

Paul Horowitz and Winfield Hill, *The Art of Electronics*, 2nd ed. Cambridge: Cambridge University Press, 1989.
Don Davis and Eugene Patronis, Jr., *Sound System Engineering,* 3rd ed. Burlington, MA: Focal Press, 2006.

This page intentionally left blank

Chapter 21

Preamplifiers and Mixers

by Bill Whitlock and Michael Pettersen

21.1 Microphone Preamplifier Fundamentals 735
  21.1.1 The Microphone as a Signal Source 735
    21.1.1.1 Electrical Model of the Microphone 735
    21.1.1.2 Interactions with Preamplifier and Cable 735
  21.1.2 Some Considerations in Practical Preamplifiers 737
    21.1.2.1 Gain and Headroom 737
    21.1.2.2 Input Impedance 737
    21.1.2.3 Noise 737
    21.1.2.4 Bandwidth and Phase Distortion 738
    21.1.2.5 Common-Mode Rejection, Phantom Power, and RF Immunity 739
21.2 Real-World Preamp and Mixer Designs 739
  21.2.1 Transformers 739
  21.2.2 Class A Circuitry 740
  21.2.3 Shure SCM268 Four-Channel Mixer 740
  21.2.4 Cooper Sound CS 104 Four-Channel ENG Mixer 740
  21.2.5 Jensen-Hardy Twin-Servo® 990 Microphone Preamplifier 741
21.3 21.3 Automatic Microphone Mixers 741
  21.3.1 The Audio Problems Caused by Multiple Open Microphones 742
    21.3.1.1 Buildup of Background Noise and Reverberation 743
    21.3.1.2 Reduced Gain Before Feedback 743
    21.3.1.3 Comb Filtering 743
    21.3.1.4 Summary 744
  21.3.2 Design Objectives for Automatic Microphone Mixers 744
    21.3.2.1 Examples of Design Objectives for an Automatic Microphone Mixer 744
  21.3.3 Controls and Features of Automatic Microphone Mixers 745
    21.3.3.1 Controls and Features Unique to Automatic Microphone Mixers 745
    21.3.3.2 External Control Capability and Status Indication of Automatic Microphone Mixers 745
    21.3.3.3 Examples of Communication Protocols Used in Automatic Microphone Mixers 746
    21.3.3.4 Number of Open Microphones Attenuation (NOMA) 746
    21.3.3.5 Restricting the Number of Open Microphones (NOM) 747
    21.3.3.6 Input Channel Attenuation 747
    21.3.3.7 Automatic Gain Control 747
  21.3.4 Types of Automatic Microphone Mixers 748
    21.3.4.1 Fixed Threshold Automatic Microphone Mixers 748
    21.3.4.2 Variable Threshold Automatic Microphone Mixers 748
    21.3.4.3 Gain-Sharing Automatic Microphone Mixers 748
    21.3.4.4 Direction-Dependent Automatic Microphone Mixers 750
    21.3.4.5 Noise-Adaptive Threshold Automatic Microphone Mixers 750
    21.3.4.6 Multivariable-Dependent Automatic Microphone Mixers 751

        21.3.4.7 Automatic Microphone Mixers with Matrix Mixing to Multiple Outputs 752
        21.3.4.8 Automatic Mixing Controller 752
        21.3.4.9 Automatic Microphone Mixers Implemented in Software 752
        21.3.4.10 Which Type of Automatic Mixer Works Best? 754
    21.3.5 Teleconferencing and Automatic Microphone Mixers 754
    21.3.6 Room and Building Acoustics 755
        21.3.6.1 Conference Room Noise 755
        21.3.6.2 Conference Room Reverberation 756
        21.3.6.3 Telephone/Transmission System Interface 757
    21.3.7 Teleconferencing Equipment 758
        21.3.7.1 Telephone Interface 758
        21.3.7.2 Microphone Considerations 759
        21.3.7.3 Microphone Mixing 759
        21.3.7.4 Loudspeaker Considerations 759
        21.3.7.5 Send Level Control 760
        21.3.7.6 Echo Canceler 760
        21.3.7.7 Historical Examples of Teleconferencing Equipment 760
        21.3.7.8 The Present and Future of Teleconferencing 762
References 763
Patent References. 763
Bibliography 763

21.1 Microphone Preamplifier Fundamentals

Microphones are transducers that typically have very low output signal levels. A voltage gain of 1000 (60 dB) or more may be required to bring the signals up to standard line levels, hence the name *preamplifier.* Amplifiers for such low-level signals are prone to problems unique to high-gain, low-noise electronic circuits. Microphone preamplifiers are available as stand-alone devices or as part of simple mixers or complex recording consoles. In this section, we will limit our discussion to preamplifiers and mixers intended for use with professional microphones that have balanced, low-impedance outputs.

21.1.1 The Microphone as a Signal Source

As discussed in Chapter 16, microphones may vary considerably in output impedance, output level or sensitivity, and self-noise. For professional microphones, impedance has a rated or *nominal* value of 150 Ω (U.S. standard) or 200 Ω (European standard). Dynamic microphones, like loudspeakers, have an *actual* impedance that varies with frequency. Note the similarity of the impedance plot of Fig. 21-2 to that of a loudspeaker. A single figure representing the impedance of such devices is usually taken as the first *minimum* that occurs after the first *maximum* as frequency is increased from a low-frequency limit. The first maximum is usually cone or diaphragm resonance. For microphones, this impedance would be measured between the signal the output pins (2 and 3 for an XLR) and is variously referred to as *output, source, internal, signal*, or *differential* impedance. Such microphones are broadly classified as *floating balanced* sources. *Floating* refers to the fact that the common-mode impedances—i.e., those from output (pins 2 and 3) to case and shield (pin 1)—are very much higher than the signal or differential impedance. As emphasized in Chapter 32, Grounding and Interfacing, balanced refers to the matching of these common-mode impedances.

21.1.1.1 Electrical Model of the Microphone

Since the Shure SM57 dynamic microphone is so popular, it will be used as an example. Fig. 21-1 shows its internal schematic, the electrical equivalent circuit of the capsule and transformer, and the combined equivalent circuit. The equivalent circuits do not model the diaphragm resonance at approximately 150 Hz. Note the pair of 17 pF capacitances to the case. These determine the common-mode output impedances that play a role in noise rejection or CMRR when the microphone is connected to a preamplifier and cable. The actual measured output impedance of an SM57 is shown in Fig. 21-2. The Shure data sheet accurately specifies the actual impedance as 310 Ω. The equivalent circuit of Fig. 21-1 models the impedance behavior above 1 kHz as well as the equivalent noise resistance. Condenser microphones generally have both lower output impedances and less variation with frequency than dynamic types.

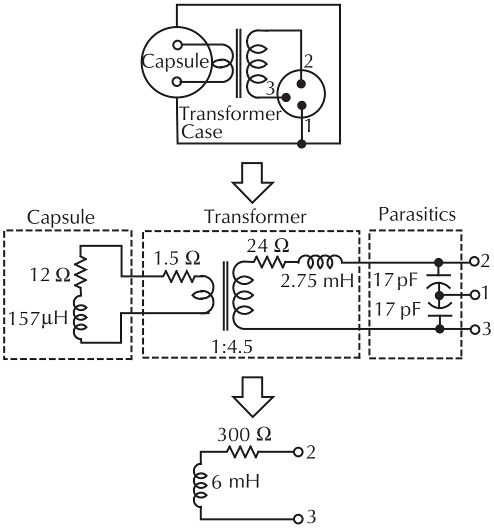


Figure 21-1. Shure SM57 schematic and equivalent circuits.

21.1.1.2 Interactions with Preamplifier and Cable

A microphone preamplifier is normally designed to recover as much of the available microphone output *voltage* as possible. Since the noise floor of the pre amplifier is nearly constant, signal-to-noise performance is improved by making its input voltage as large as possible. It is very important to understand that the fraction of available microphone voltage actually delivered to the preamplifier depends on both the output impedance of the microphone and the input impedance of the preamplifier.

As shown in Fig. 21-3, these two impedances effectively form a voltage divider. The voltage lost in the output impedance Z_S of the microphone depends on the input impedance Z_L of the preamp. *Loading loss*, usually expressed in dB, compares the output voltage with some specified load to the output voltage under open circuit or unloaded conditions. For example, a 150 Ω impedance (actual) microphone will deliver 91% of its unloaded voltage when loaded by a preamplifier having a 1.5 kΩ input impedance. The loading loss is

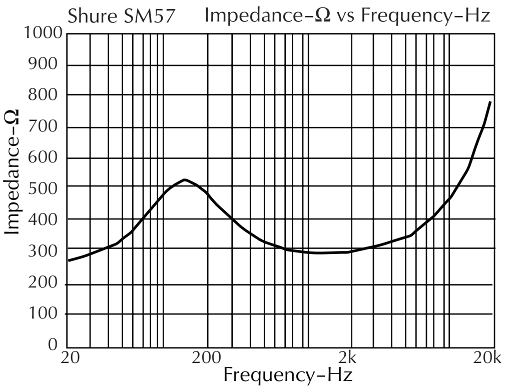
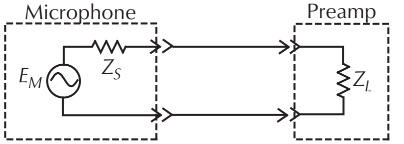


Figure 21-2. Measured output impedance of the Shure SM57.



Input voltage at preamp $E_P = E_M \times Z_L/Z_S + Z_L$

Figure 21-3. Microphone and preamplifier voltage divider.

then 20 × log 0.91, which is 0.8 dB. Generally, loading loss is negligible (under 1 dB) if load impedance is ten or more times the source impedance. Therefore, as discussed in Chapter 16, it is neither desirable nor necessary to "match" the impedances of the preamplifier and microphone. If impedances are matched, half the available output voltage from the microphone is lost, degrading signal-to-noise ratio by 6 dB. Although impedance matching transfers maximum power, this is *not* what we want.

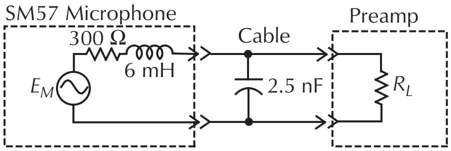


Figure 21-4. Low-pass filter formed by microphone, cable, and preamplifier.

When a microphone is connected to a cable and a preamplifier, a passive two-pole (12 dB/octave) low-pass LC filter is formed as shown in Fig. 21-4. The behavior of LC filters as they approach their cutoff or resonant frequency is controlled by resistive elements in the filter. This resistive damping is largely provided by the input resistance R_L of the preamplifier. Fig. 21-5 shows the deviation in frequency response of a Shure SM57 microphone loaded by the 2.5 nF capacitance of 150 ft of typical microphone cable and three different values of preamplifier input resistance. The upper curves, 10 kΩ and 3 kΩ are typical of preamps that don't use an input transformer. Note the high-frequency response peaking caused by insufficient damping. The lower curve, 1.5 kΩ, is typical of a preamp using an input transformer.

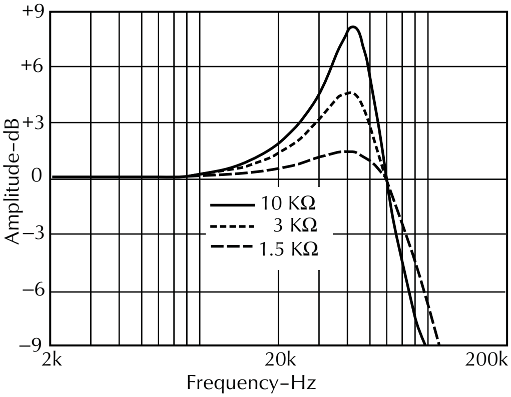


Figure 21-5. Effect of load resistance on damping.

Capacitance of shielded twisted pair cable is usually specified as that from one conductor to the other conductor and the shield. Belden 8451, for example, is listed at 34 pF/ft. However, the differential signal is affected by the capacitance between the conductors, which is about half that, or 17 pF/ft. With dynamic microphones, high cable capacitance causes high-frequency roll-off. For the SM57 microphone, about a thousand feet of this cable (about 17 nF) will limit high-frequency bandwidth to about 15 kHz. Because condenser microphones use internal amplifiers to drive the output cable, high cable capacitance can cause distortion. If the amplifier has limited output current, it will distort or clip high-level, high-frequency (i.e., high slew rate) signals such as vocal sibilance or a cymbal crash. "Star-quad" cable, although it offers amazing freedom from magnetic pickup problems, has about twice the capacitance per foot of standard cable. This fact must be seriously considered for long cables.

Keep in mind, however, that other types (or even models) of microphones may behave quite differently, depending on their exact equivalent circuit. For example, some condenser types have low (around 30 Ω) and almost purely resistive output impedances while some dynamic types can have actual midband impedances over 600 Ω.

To a greater or lesser degree, the frequency response of any microphone will be affected by the load capacitance of the connected cable and preamplifier as well as the input impedance characteristics of the preamplifier. Perhaps this is why the selection of microphones and preamplifiers is such a subjective issue.

21.1.2 Some Considerations in Practical Preamplifiers

Because many aspects of preamplifier circuit design and the tradeoffs involved are discussed in Chapter 25, Sections 25.6 and 25.9, we will discuss only a few topics here.

21.1.2.1 Gain and Headroom

Microphone preamplifiers commonly have maximum voltage gains of about 60 dB to 80 dB and minimum gains from 0 dB to 12 dB. A typical microphone, such as the Shure SM57, will have an output of 1.9 mV or –52 dBu with a 94 dB SPL acoustic input. For a very high acoustic input of 134 dB SPL, its output would be 190 mV or –12 dBu. But a high-sensitivity microphone such as the Sennheiser MKH-40 will have an output of 25 mV or –30 dBu at 9 dB SPL and 2.5 V or +10 dBu at 134 dB SPL. Such high input levels can actually require the preamplifier to have a loss (i.e., negative gain) to produce usable line level output. Such high input levels can also overload the preamp. Both problems are most commonly avoided with an input attenuator or pad, typically of 20 dB. See Chapter 11 for a discussion of the distortion and level handling characteristics of audio transformers.

21.1.2.2 Input Impedance

As shown in Fig. 21-5, some input transformers have input impedances that load the microphone and, as discussed in the preceding section, alter the response of the system at frequency extremes. However, well-designed transformers such as the Jensen JT-16B have substantially flat input impedance as shown in Fig. 21-6.

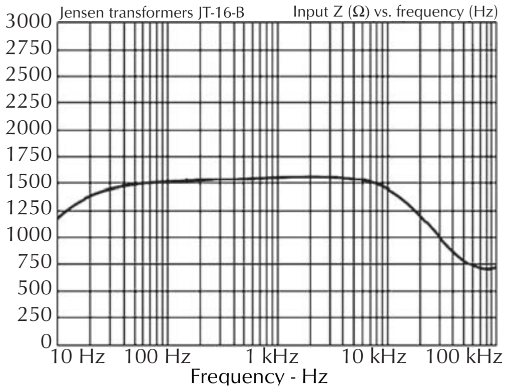


Figure 21-6. Input impedance of a Jensen JT-16B input transformer.

21.1.2.3 Noise

The random motion of electrons in electrical conductors creates a voltage variously called *thermal noise, white noise*, or *Johnson noise* after its first observation by J. B. Johnson of Bell Labs in 1927. Thermal noise voltage is proportional to both temperature and the resistance of the conductor and is calculated as follows:[1]

$$E_t = \sqrt{4kTR\Delta f} \tag{21-1}$$

where,

E_t is the thermal noise in rms volts,
k is Boltzmann's constant or 1.38×10^{-23} Ws/°K,
T is the temperature of the conductor in degrees Kelvin,
R is the resistance of the conductor in ohms,
Δf is the noise bandwidth in Hertz.

At a room temperature of 300°K (80°F or 27°C), $4kT = 1.66 \times 10^{-20}$. For noise in the audio band of 20 Hz to 20 kHz, bandwidth is 19.98 kHz. It's important to note that noise bandwidth here refers to a rectangular "brick wall" response, not the more conventional measure at the –3 dB points. For a 150 Ω resistance under these conditions, noise is

$$\begin{aligned} 223\ \text{nVrms} &= -133.0\ \text{dBV} \\ &= -130.8\ \text{dBu}. \end{aligned}$$

For a 200 Ω resistance under the same conditions, noise is

$$\begin{aligned} 258\ \text{nVrms} &= -131.8\ \text{dBV} \\ &= -129.5\ \text{dBu}. \end{aligned}$$

Here we use the nominal impedance of an idealized microphone simply to allow a simple but fair comparison of preamplifier noise performance.

Regardless of whether the conductor is copper wire, silver wire, an expensive metal-film resistor, or a cheap carbon resistor, the thermal noise is exactly the same! *Excess* noise refers to additional noise generated when dc flows in the resistor. Excess noise varies markedly with resistor material and construction. Note that only the resistive portion of impedance generates noise—pure inductors and capacitors do not generate thermal noise. Therefore, in our Shure SM57 circuit model of Fig. 21-1, thermal noise is generated by the 300 Ω resistance but not by the 6 mH inductance.

In a practical microphone preamp, we are usually concerned with the signal-to-noise ratio at the output. Although there may be many sources of internal noise in the preamplifier and its gain may be varied over a wide range, for simplicity noise is usually stated in terms of *EIN* or *equivalent input noise*. This simplification works because, in a good design, the dominant noise source is the first amplification stage and subsequent stages contribute no significant noise.

As shown in Fig. 21-7, the EIN has three components:

1. E_t: the thermal noise of the source resistance.
2. E_n:the voltage noise of the amplifier.
3. I_n:the current noise of the amplifier.

When noise voltages are produced independently and there is no relationship between their instantaneous amplitudes or phases, they are said to be uncorrelated. Total noise power is the sum of individual noise powers. Therefore, the resultant voltage is the square root of the sum of the squares of the individual voltages. For example, adding two uncorrelated 1 V noises will result in a 1.4 V noise because

$$\begin{aligned} E &= \sqrt{1^2 + 1^2} \\ &= \sqrt{2} \\ &= 1.414. \end{aligned} \tag{21-2}$$

When adding two noises, unless the second is a third or more of the first (less than 10 dB difference), it will have little effect on the total. For any amplifier, minimum total noise is added when the source resistance is such that I_N flowing through the source creates a noise voltage equal to E_N. This source resistance is called the *optimum* source resistance for that particular amplifier. Perhaps the most useful function of an input transformer in a microphone preamplifier is to convert, as explained in Chapter 11 Audio Transformers, the impedance of the microphone to this optimum value in order to maximize SNR.

Measurement of noise is fertile ground for technical misrepresentation. Some rather unbelievable EIN numbers have appeared over the years. Most were based on measurements taken with the preamplifier input shorted, which ignores the noise contributions of both R_S (source resistance) and I_N (amplifier current noise), leaving only E_N (amplifier voltage noise). Bias current noise generates additional voltage noise when it flows in the source impedance (not just resistance). In this case the inductance of our SM57 model will indirectly contribute real-world noise. To have any meaning at all, EIN must specify the source impedance. With a 150 Ω source resistance,

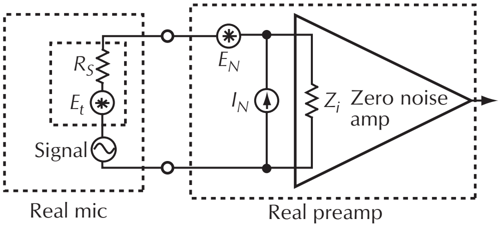


Figure 21-7. Contributions to equivalent input noise.

$$\begin{aligned} EIN &= 223 \text{ nVrms} \\ &= -133.0 \text{ dBV} \\ &= -130.8 \text{ dBu} \end{aligned}$$

for an ideal noiseless amplifier. If the preamplifier noise is equal to that of the source, EIN will be 3 dB higher or –130.0 dBV = –127.8 dBu. *Noise figure*, or NF, is a measure of SNR degradation attributed to the amplifier—in this case 3 dB. From an engineering point of view there is little point in attempting to achieve NF below 3 dB.[2]

Note that the thermal noise Eq. 21-1 also includes a term for bandwidth. Noise specifications such as EIN frequently appear in data sheets without a specified noise bandwidth. All other things equal, noise increases as the square root of bandwidth. Therefore, there is 1.25 dB less noise in a 15 kHz bandwidth, and 3 dB less noise in a 10 kHz bandwidth, than in a 20 kHz bandwidth. Likewise, while measurements such as A-weighted noise are both legitimate and useful, they cannot be directly compared to unweighted measurements. When comparing noise specifications, be sure it's an "apples to apples" comparison.

21.1.2.4 Bandwidth and Phase Distortion

Performance in the *time domain*, or waveform fidelity, is critically important to accurate music reproduction. Accurate time domain performance, sometimes called *transient response*, requires low phase distortion. Pure time delays exhibit a linear phase versus frequency characteristic. True phase distortions are expressed as DLP or *deviations* from this linear phase relationship. Phase *shift* is not necessarily phase distortion.[3] In order to achieve a DLP of 5° or less from 20 Hz to 20 kHz, frequency response must extend from 0.87 Hz to 35 kHz, assuming 6 dB per octave (first order) filter responses. If the low-pass filter is a second order Bessel, the cutoff frequency can be as low as 25 kHz.[4] Notice that extreme high-frequency response is not required, but extended low-frequency response is!

Phase distortion not only alters musical timbre, but it has potentially serious system headroom implications as well. Even though frequency response may be flat, peak signal *amplitudes* can increase up to 15 dB after passing through a network with high phase distortion. This can be a serious problem in digital recording systems. Even ultrasonic phase distortions caused by undamped resonances can excite complex *audible* cross-modulation products in subsequent nonlinear (any real world) amplifier stages.[5]

Low-frequency phase distortions are often described as muddy bass and high-frequency phase distortions as harshness or midrange smear. The complex cross-modulation products are usually described as dirty sounding and often are the cause of listener fatigue.

21.1.2.5 Common-Mode Rejection, Phantom Power, and RF Immunity

Common-mode rejection, as discussed in Chapter 11, Audio Transformers, is not just a function of the amplifier input circuitry. It depends on the impedance balance achieved by the combination of the microphone's output circuitry, cable, and the preamp's input circuitry. Common-mode rejection ratio (CMRR) is seldom an issue with dynamic microphones because, as shown in Fig. 21-1, the common-mode impedances are small parasitic capacitances. However, when phantom power is involved, very high CMRR can be difficult to achieve. The circuitry in the microphone that extracts phantom power from the two signal lines, as shown in the examples in Chapter 16, Microphones, can unbalance their line impedances to ground. The resistors that supply phantom power, shown in the preamplifier of Fig. 21-13, must also be tightly matched to achieve high CMRR. For example, CMRR may be limited to 93 dB if ±0.1% resistors are used and may be limited to 73 dB if ±1% resistors are used. For comparison, the JT-16B transformer used in Fig. 21-13 achieves a CMRR of 117 dB when phantom power resistors are absent. Sometimes, as in Fig. 21-9, phantom power is supplied through a center-tap on a microphone input transformer. This presents a transformer design problem that can be even more difficult—simultaneously matching both the number of turns and the dc winding resistance on each side of the center tap.

RF interference, usually in the form of common-mode voltage, is another potential problem for microphone preamplifiers because it is likely to be demodulated in amplifier circuitry. In transformer-less circuits, suppression measures usually consist of capacitors from each input to ground and sometimes series resistors, chokes, or ferrite beads. Unless the capacitors are carefully matched, they will *unbalance* the common-mode input impedances and degrade CMRR. Because they also lower common-mode input impedances, they can make the circuit more sensitive to normal impedance imbalances in the microphone. These tradeoffs can be largely avoided by using a Faraday-shielded input transformer that has inherent RF suppression characteristics.

A good microphone preamplifier should also be free of the so-called pin 1 problem. The microphone cable should be free of shield-current-induced noise (SCIN), which can be a serious problem with foil shield and drain wire construction. Both of these problems are discussed in Chapter 15.

21.2 Real-World Preamp and Mixer Designs

21.2.1 Transformers

Manufacturers of microphone preamplifiers have a natural desire to differentiate their product from all others. One of the major divisive issues is the use of audio transformers. According to the antitransformer camp, all audio transformers have inherent limitations such as limited bandwidth, high distortion, mediocre transient response, and excessive phase distortion. Unfortunately, many such transformers do exist and not all of them are cheap. The makers of such transformers are simply ignorant of sonic clarity issues, have a poor understanding of the engineering tradeoffs involved, or are willing to take manufacturing shortcuts that compromise performance to meet a price.

As stated earlier, bandwidth and phase distortion are intimately linked in any electronic device. A very high level of performance can be reached with proper transformer design. Consider the Jensen JT-16B microphone input transformer. Its frequency response is −3 dB at 0.45 Hz and 220 kHz and −0.06 dB from 20 Hz–20 kHz, with a second order Bessel high-frequency roll-off characteristic. Low frequency roll-off is less than 6 dB per octave owing to properties of the core material, which further improves phase performance. Its deviation from linear phase is under 2° from

20 Hz–20 kHz, giving it truly excellent waveform fidelity and square-wave response.

As discussed in detail in Chapter 11, Audio Transformers, audio transformer distortion is quite different from electronicdistortion in ways that make it unusually benign. First, transformer distortion is frequency and level dependent. Significant distortion occurs only at low frequencies and high signal levels, typically dropping to under 0.001% above a few hundred hertz. Second, the distortion is nearly pure third harmonic and is *not* accompanied by the high levels of much more irritating intermodulation distortion that occurs in electronics.

A high degree of RF attenuation, both normal mode and commonmode, is also inherent in transformers that contain Faraday shields. For example, in Jensen designs, common-mode attenuation is typically over 30 dB from 200 kHz–10 MHz. And, as discussed in Chapter 16, transformers enjoy a great CMRR advantage over most electronically balanced input stages because they are relatively insensitive to the impedance imbalances that normally exist in real-world signal sources. If well designed and properly applied, audio transformers qualify as true high-fidelity devices. They are passive, robust, and stable and have significant advantages, especially in electrically hostile environments.

21.2.2 Class A Circuitry

Another divisive issue among preamplifier manufacturers involves class A circuitry. Although it has certain advantages, it is not necessarily inherently superior to much more widely used class AB designs. Class A operation occurs when the active device (or devices in the case of a push-pull output stage) conducts current during the entire 360° signal cycle. Class AB occurs when each device conducts for more than 180° but less than 360°. In class B operation, each device conducts for exactly 180°. In class C, conduction is less than 180° and this is generally done only in RF circuits or where intentional distortion is desired.

Most op-amp output stages operate class AB to avoid crossover distortion of small signals. Practical active devices are generally unable to behave linearly near zero current (cutoff) as is required for low-distortion pure class B operation. A small idling or quiescent current flows in both devices at zero signal and operation remains class A (both devices conducting for full signal cycle) up to some signal level, at which point one device begins to be cut off for part of the cycle, producing class AB operation.

For example, in the Jensen-Hardy 990 amplifier module used in the circuit of Fig. 21-13, this output stage quiescent current is about 15 mA. Therefore, amplifier operation is class A until peak output current (plus or minus) reaches about 15 mA. Peak output current, of course, depends on peak signal level and load impedance. For example, the output voltage clips at about 24 Vpeak, so any load impedance higher than about 1.6 kΩ results in class A operation at all times. Likewise, with a 600 Ω load, operation is class A until output signal level reaches ±9 Vpeak. Above that peak level, operation becomes class AB. The "front end" circuitry of the 990, like most operational amplifiers, always operates class A unless the output is clipped. The line between class A and AB operation is very distinct: operation is no longer class A as soon as current in any active device (vacuum tube or transistor) becomes zero. The main advantage of class A circuit designs is that the curvature of the nonlinearity plot is likely to be smoother (i.e., free of a sharp discontinuity at crossover) so that there will be fewer problems related to negative feedback, slew rate, and gain-bandwidth limitations.

21.2.3 Shure SCM268 Four-Channel Mixer

The Shure SCM268 is an example of a compact, simple mixer with basic features. Notable features include transformers on balanced inputs and outputs, mic or line level output, phantom power, and optional pads allowing for balanced line level inputs, Fig. 21-8. A functional block diagram is shown in Fig. 21-9.

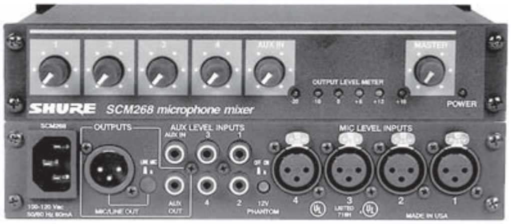


Figure 21-8. Shure SCM268 four-channel microphone mixer. Courtesy Shure Incorporated.

21.2.4 Cooper Sound CS 104 Four-Channel ENG Mixer

The Cooper Sound Systems CS 104 is an example of a portable, battery-powered mixer with a number of sophisticated features, Fig. 21-10. Notable features include stereo mixing, pan pots and channel linking,

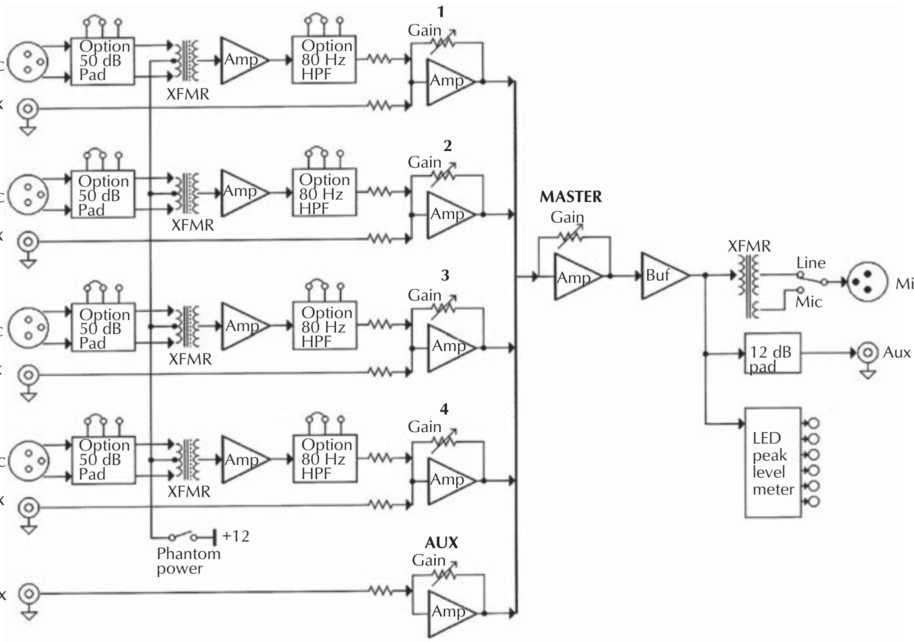

Mic
1
aux
Option
50 dB
Pad
XFMR
Amp
Option
80 Hz
HPF
1
Gain
Amp
Mic
2
aux
2
MASTER
Gain
Mic
3
aux
3
Buf
XFMR
Line
Mic
Mic/Line
12 dB
pad
Aux
Mic
4
aux
4
LED
peak
level
meter
+12
Phantom
power
AUX
Gain
Aux

Figure 21-9. Block diagram of a Shure SCM268.

transformers on main inputs and outputs, built-in stereo limiter, input overload indicators, selectable high-pass filters, prefade listen, tape monitor, and built-in tone and slate functions. Fig. 21-12 is its functional block diagram.

21.2.5 Jensen-Hardy Twin-Servo® 990 Microphone Preamplifier

The Jensen-Hardy Twin-Servo® 990 Microphone Preamplifier is an example of a high-performance design, Fig. 21-11. It features patented discrete component 990 amplifier modules that combine low-input noise, high-output voltage and current, low distortion, and high gain-bandwidth performance that is unavailable with integrated circuits. It uses two cascaded variable gain stages per channel to maintain high bandwidth and low distortion overall. Unlike most designs, this topology also keeps EIN very low at the lowest gain settings. Extended low-frequency response is preserved by using dc servo feedback circuitry to eliminate coupling capacitors and their attendant problems, Fig. 21-13.

MASTER
CSS
CSS
COOPER SOUND SYSTEMS

Figure 21-10. Cooper Sound CS 104 battery-powered portable ENG/EFP four channel audio mixer.

jensen
TWIN SERVO

Figure 21-11. Jensen Twin-Servo® 990 four channel microphone preamplifier. Courtesy Jensen Transformers, Inc.

21.3 21.3 Automatic Microphone Mixers

Automatic microphone mixers, also known as *voice-activated mixers* or *sound-activated mixers*, have become a necessary part of sound systems designed for speech. All automatic microphone mixers have a fundamental function: to attenuate (reduce in level) any microphone that is not being spoken into by a talker, and conversely, to rapidly activate any microphone that is being spoken into by a talker. An automatic microphone mixer should

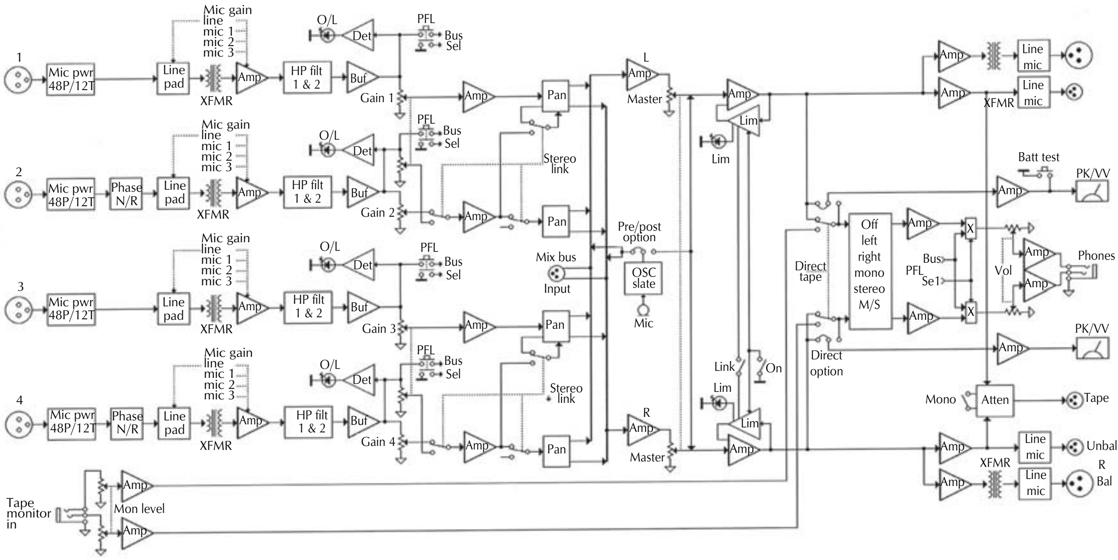

Mic gain
line
mic 1
mic 2
mic 3
1
2
3
4
Mic pwr 48P/12T
Phase N/R
Line pad
XFMR
Amp
HP filt 1 & 2
Buf
Gain 1
Gain 2
Gain 3
Gain 4
O/L
Det
PFL
Bus
Sel
Pan
Stereo link
L
R
Master
Lim
Link
On
Mix bus
Input
Pre/post option
OSC slate
Mic
Direct tape
Direct option
Off left right mono stereo M/S
Bus
Se1
Vol
Phones
Batt test
PK/VV
Mono
Atten
Tape
Line mic
Unbal
Bal
Tape monitor in
Mon level

Figure 21-12. Block diagram of the Cooper Sound CS 104.

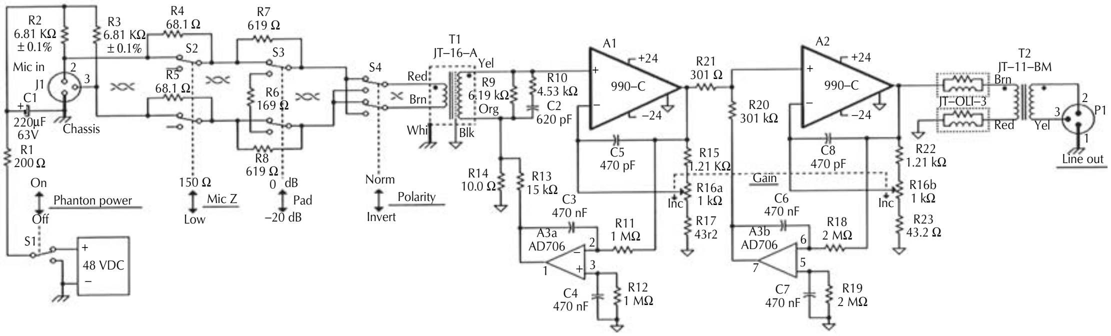

R2 6.81 KΩ ±0.1%
R3 6.81 KΩ ±0.1%
R4 68.1 Ω
R5 68.1 Ω
R6 169 Ω
R7 619 Ω
R8 619 Ω
S2
S3
S4
Mic in
J1
C1 220μF 63V
Chassis
R1 200 Ω
On
Phanton power
Off
S1
48 VDC
150 Ω
Mic Z
Low
0 dB
Pad
−20 dB
Norm
Polarity
Invert
T1 JT–16–A
Red
Brn
Whi
Yel
Org
Blk
R9 6.19 kΩ
R10 4.53 kΩ
C2 620 pF
R14 10.0 Ω
R13 15 kΩ
C3 470 nF
A1
990–C
+24
−24
C5 470 pF
A3a AD706
C4 470 nF
R11 1 MΩ
R12 1 MΩ
R21 301 Ω
R15 1.21 KΩ
R16a 1 kΩ
Inc
R17 43r2
R20 301 kΩ
A2
990–C
Gain
C8 470 pF
C6 470 nF
A3b AD706
C7 470 nF
R18 2 MΩ
R19 2 MΩ
R22 1.21 kΩ
R16b 1 kΩ
R23 43.2 Ω
JT–OLI–3
T2 JT–11–BM
Brn
Red
Yel
P1
Line out

Figure 21-13. Signal path schematic of Jensen Twin-Servo 990 microphone preamplifier. Courtesy Jensen Transformers, Inc.

be considered when the number of microphones required for the sound system is four or greater.

When used in a sound reinforcement system, an automatic microphone mixer provides a significant increase in gain before feedback when multiple microphones must be used without a sound engineer. It also improves the quality of the sound system output by reducing the amount of extraneous room sound being picked up, and by reducing comb filtering. In addition, it automatically adjusts system gain to compensate for the number of microphones in use at any instant. Thus, an automatic microphone mixer attempts to provide the same system control that might be produced by a human sound engineer.

As automatic microphone mixers are optimized for speech applications, their use in musical applications is not recommended. Mixing microphones for music is as much art as science, and therefore the artistic judgment of a human sound engineer is much preferred to the electronic decision process of an automatic microphone mixer.

In summary, when used in a speech sound system with multiple microphones, the ideal automatic microphone mixer assures that the number of active microphones at any moment equals the number of active talkers at the same moment. All unused microphones at that moment are attenuated.

21.3.1 The Audio Problems Caused by Multiple Open Microphones

High-quality audio becomes progressively more difficult to achieve as the number of open microphones increases. All audio systems face the same problems whenever multiple open microphones are needed. These problems are:

1. Build-up of background noise and reverberation.
2. Reduced gain before feedback.
3. Comb filtering.

These problems can plague boardrooms, city council chambers, conference centers, houses of worship, teleconferencing rooms, radio talk shows—anywhere multiple microphones are used. Since audio quality rapidly deteriorates as the number of open microphones increases, the solution is to keep the minimum number of microphones open that will handle the audio. An automatic microphone mixer keeps all unused microphone input channels attenuated, and activates any microphone spoken into within milliseconds.

21.3.1.1 Buildup of Background Noise and Reverberation

The first problem of multiple open microphones is the buildup of background noise and reverberation. This buildup can adversely affect the quality of recordings or broadcasts originating from the audio system. Consider the case of a city council with eight members and eight microphones. For this example, only one member is talking. If all eight microphones are open when only one microphone is needed, the audio output will contain the background noise and reverberation of all eight microphones. This means the audio signal will contain substantially more background noise and reverberation than if only the talker's microphone were open. This buildup of background noise and reverberation greatly deteriorates the audio quality. Speech clarity and intelligibility always suffer as background noise and reverberation increase.

As the number of open microphones increases, the background noise and reverberation in the audio output also increase. In our city council example, the audio output from eight open microphones would contain 9 dB more background noise and reverberation than a single open microphone. To the human ear, the noise would sound almost twice as loud when all eight microphones were open.

To minimize background noise and reverberation buildup, an automatic microphone mixer activates only the microphone(s) being addressed and employs a NOMA circuit. NOMA is an acronym for number of open microphones attenuator. NOMA systematically decreases the master gainwhenever the number of open microphones increases. Without NOMA, the audio system would produce objectionable noise modulation (pumping and breathing) as background noise and reverberation increase and decrease with the number of open microphones. With a properly designed automatic microphone mixer, background noise and reverberation remain constant no matter how many or few microphones are activated.

21.3.1.2 Reduced Gain Before Feedback

The second problem of multiple open microphones is reduced gain before feedback. Acoustic feedback ("howling") can be a problem anytime a sound reinforcement (PA) system is used. To avoid feedback, PA systems are operated below the point where the system becomes unstable and starts to howl. However, this feedback safety margin is reduced each time another microphone is opened. Have one too many open microphones and the result is feedback.

The automatic microphone mixer solution is to keep unused microphones turned off and utilize NOMA. As more microphones are activated, the overall gain will remain constant thanks to the NOMA circuit. An automatic microphone mixer assures that if the audio system does not feedback when any one microphone is open, the system will remain feedback free even if all the microphones are open.

21.3.1.3 Comb Filtering

The third problem of multiple open microphones is comb filtering. Comb filtering occurs when open microphones at different distances from a talker are mixed together, Fig. 21-14. Since sound travels at a finite speed, the talker's voice arrives at the microphones at different times. When combined in a mixer, these out-of-step microphone signals produce a combined frequency response very different from the frequency response of a single microphone. (A frequency response chart of the out-of-step signals looks like the teeth of a hair comb, thus the name.) The aural result of comb filtering is an audio signal that sounds hollow, diffuse, and thin.

The solution to comb filtering also is keeping the number of open microphones to an absolute minimum. By automatically turning off unused microphones, an automatic microphone mixer reduces comb filtering and the resultant poor audio.

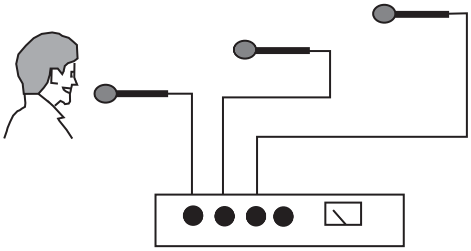


Figure 21-14. Comb filtering occurs whenever open microphones at different distances from a talker are mixed together.

21.3.1.4 Summary

1. Keeping the number of open microphones to a minimum always improves overall audio quality.
2. The primary function of an automatic microphone mixer is to keep unused microphone input channels attenuated (turned down or off) and to instantaneously activate microphones when needed.
3. Buildup of background and reverberant noise, reduced gain before feedback, and comb filtering can all be controlled by using an automatic microphone mixer.

21.3.2 Design Objectives for Automatic Microphone Mixers

As shown in Fig. 21-15, a conventional microphone mixer in a sound system amplifies the signal from each microphone and combines these amplified signals together to produce a single output. This output feeds a power amplifier and then one or more loudspeakers. Each doubling of the number of open microphones feeding into a sound system reduces the available gain before feedback by 3 dB. This fact surprises the layman who often believes that more microphones equate to the sound system being louder, not softer. A sound system with numerous microphones easily becomes ineffective if a sound engineer is not present to control levels and switch off unused microphones. Since gain before feedback can often be marginal because of the acoustical characteristics of a room, an automatic microphone mixer may be the only way to provide adequately loud program levels to the audience with an unattended sound system.

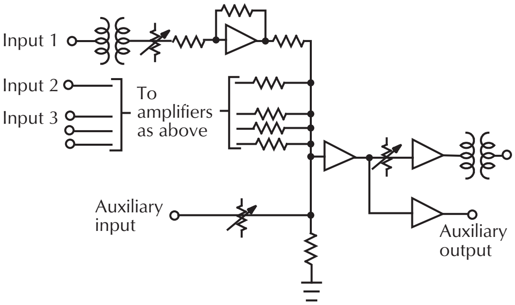


Figure 21-15. Simplified diagram of a microphone mixer.

21.3.2.1 Examples of Design Objectives for an Automatic Microphone Mixer

1. Keeps the sound system gain below the threshold of feedback instability.
2. Requires no operator or sound technician at the controls.
3. Does not introduce spurious, undesirable noise or distortion of the program signals.
4. Can be installed as easily as a conventional mixer.
5. Responds only to the desired speech input signals and is relatively unaffected by extraneous background noise signals.
6. Activates input channels fast enough that no audible loss of speech signals occurs.
7. Allows more than one talker on the system when required by the discussion content while still maintaining control of the overall sound system gain.
8. Adjusts the system gain to compensate for a range of talker input levels.
9. Provides system status outputs for peripheral equipment control and can interface with external control systems for advanced system design if required.

The automatic microphone mixer operation should provide relatively easy and very rapid input activation. Desired speech from a talker should cause immediate activation of the appropriate input channel, which may not always happen if the design of an automatic microphone mixer is poor. Also, random false activation of microphones remote from the talker can occur with some automatic microphone mixer designs. However, this false activation is typically not troublesome as the false signals are normally much lower in level than the desired talker signal. The automatic microphone mixer is doing its job if all talkers are clearly heard by the audience when they speak, and the sound system remains below the point of feedback.

An automatic microphone mixer cannot improve the performance of microphones. Its primary benefit comes from limiting the number of microphone signals fed to the mixer output. A side benefit is often the apparent increase of critical distance in a multiple microphone system. (Critical distance is defined as a point in the room where the direct signal of the talker equals the reflected signal of the talker, i.e., 50% direct signal and 50% reverberant signal.) Because unused microphones remote from the talker are attenuated, room reverberation and ambience that would otherwise be amplified are reduced.

21.3.3 Controls and Features of Automatic Microphone Mixers

Automatic microphone mixers have many of the same controls and features of manual microphone mixers. Examples are:

- Level control for each input channel.
- Master level control for each output channel.
- Input signal attenuation (“trim”).
- Phantom power.
- Two or three band equalization for each input channel.
- Output level metering.
- Output signal level limiter.
- Nonautomatic auxiliary inputs.
- Headphone output with level control.

These controls and features may be configured in hardware–e.g., switches, potentiometers, LED strings—or they may be configured in software. In either case, the function of the control or feature remains the same.

21.3.3.1 Controls and Features Unique to Automatic Microphone Mixers

As automatic microphone mixers typically perform more functions than a manual microphone mixer, there are controls and features that are unique to automatic microphone mixers.

Input Channel Threshold. Determines at what signal level a gated automatic microphone mixer input passes the incoming microphone signal to the mixer’s output.

Input Channel On Indicator. Illuminates to indicate that an input channel is passing the microphone signal onto the mixer output.

Direct Output for Each Input Channel. Provides an isolated output for each input channel that is unaffected by the automatic microphone mixer action.

Last Microphone Lock On. Keeps on the most recently activated input channel on a gated automatic microphone mixer until another input channel is activated. This maintains room ambience when the automatic microphone mixer is used to provide a broadcast feed, a recording feed, or a feed to an assistive hearing system.

Hold Time. Keeps an activated input channel on a gated automatic microphone mixeron for a period of time after speech has ceased. This feature bridges the natural gaps that occur in speech patterns.

Input Attenuation. Determines how much gain reduction is applied to an input channel of a gated automatic microphone mixer when the channel is not activated. Typical range of adjustment is 3 dB to 70 dB of attenuation, with 15 dB being a common value.

Decay Time. Establishes the time required for an input of a gated automatic microphone mixer to be lowered from the activated state to the attenuated state. Decay time is always in addition to the hold time.

Manual/Auto Select. Allows the automatic microphone mixer to operate in a nonautomatic (manual) mode.

21.3.3.2 External Control Capability and Status Indication of Automatic Microphone Mixers

Most automatic microphone mixers include the ability to be controlled by external switches, potentiometers, touch screens, personal computers, and other types of control devices. These devices are connected to the automatic microphone mixer via screw terminals or multipin connectors on the mixer’s rear panel. The controllable functions and the communication protocol depends upon the manufacturer and model of the automatic microphone mixer. Examples of automatic microphone mixer functions that can be externally controlled follow.

Gain of an Input Channel or the Master O\output. In a courtroom, the court clerk could control the volume level of the witness microphone or the entire sound system using a potentiometer located at a distance from the automatic microphone mixer.

Mute an Input Channel. In a city council chamber, a council member could have a privacy or “cough” switch located near the microphone.

Global Mute of All Input Channels. In a government hearing room, the presiding member could mute all inputs to regain control of a meeting.

Permanent Activation of an Input Channel. In a house of worship, a microphone over the congregation could be kept on at all times to provide constant room ambience for a hearing assistance system.

Routing of Input Channels to Different Outputs. In a hotel meeting facility with movable dividing walls, input channels could be sent to different banks of loudspeakers depending on the room configuration.

Status Terminal. As input channels are activated and attenuated by the automatic mixing process, it is valuable to have a status terminal that indicates if a particular input channel is activated or not. This status terminal, also known as a *gate terminal*, can be thought of as an electronic switch that changes from open to closed based on the activity of the input channel. Examples of use for a status terminal:

- **Control of an LED or lamp to indicate input channel activity.** In a city council chamber, a council member could have a tally light located near the microphone indicating when the microphone's input channel is activated by the automatic microphone mixer.
- **Control of a relay used to attenuate the nearest loudspeaker.** As the typical feedback path in a sound system is between a microphone and the nearest loudspeaker, attenuating the closest loudspeaker when a microphone is active could improve gain before feedback.
- **Control of a video switcher connected to multiple cameras.** In a courtroom, the proceedings could be videotaped by using cameras that follow the activation of input channels by the automatic microphone mixer.
- **Mute other input channels.** In a hotel meeting facility, one input channel could override all others in case of an emergency announcement.

Combining the externally controlled functions with the status terminals provides hundreds of unique system configurations. Most manufacturers of automatic microphone mixers have documentation of such configurations, often printed in product installation manuals and available on the manufacturer's web site. As previously noted, the communication protocol used to interpret the status terminals and control the mixer functions depends upon the manufacturer and model of the automatic microphone mixer.

21.3.3.3 Examples of Communication Protocols Used in Automatic Microphone Mixers

Contact Closure Protocol. The most basic of communication protocols, contact closure is provided by a simple single pole/single throw (SPST) switch or relay. The switch is connected to two terminals on the mixer that control a certain function —e.g., mute of an input channel. When the switch is closed, the input channel is muted. When the switch is open, the input channel is unmuted and can be activated.

Resistance Change or Voltage Change Protocol. Used primarily to control signal levels via a VCA (voltage-controlled amplifier), this protocol requires that defined changes in resistance or voltage be applied to the mixer's control terminals. In response, the VCA in the automatic microphone mixer will change the level of the audio signal.

TTL (Transistor-Transistor Logic). An electronic protocol established in the 1960s, TTL is simple to use. A control terminal on the automatic microphone mixer has one of two states: logic high (+5 Vdc) or logic low (0 Vdc). A status terminal could be logic high when a mixer input channel is attenuated and logic low when the mixer input channel is activated. This change of voltage informs an external control device that there is a change in the input channel status and some predetermined action should take place —e.g., illuminate an LED or switch on a camera.

RS-232. Used for communication with a computer, RS-232 is another common electronic protocol. RS-232 is most often used when proprietary control software is supplied with the automatic microphone mixer or when the mixer is connected to a control system such as those manufactured by Crestron or AMX.

RS-422. Basically a balanced line version of RS-232, RS-422 is designed for situations where an extremely long cable run must be used to connect the automatic microphone mixer to the external control device.

21.3.3.4 Number of Open Microphones Attenuation (NOMA)

NOMA is a function shared by all well-designed automatic microphone mixers. It is a simple method of ensuring system stability by automatically reducing the mixer output gain in proportion to the number of activated input channels. NOMA offsets the increase of gain that occurs as more microphones are activated. The attenuation in decibels should vary as:

$$\text{Attenuation in dB} = 10\log N \qquad (21\text{-}3)$$

where,

N is the number of activated microphones.

While NOMA helps maintain stable gain in a sound system as the number of activated microphones varies, it does not limit the number of microphones that can be activated.

21.3.3.5 Restricting the Number of Open Microphones (NOM)

Recent developments in automatic microphone mixer design have led to a feature best described as a *NOM restrictor*. This feature restricts the number of active input channels to a predetermined amount. For example, in a large legislative system with 100 microphones, it makes little sense to allow all 100 microphones to be active at any instant, even if all 100 legislators are talking.

Restricting the NOM to 5 microphones of the 100 allows spirited debate while not subjecting the audience to the cacophony of 100 open microphones.

21.3.3.6 Input Channel Attenuation

Gating automatic microphone mixers use some form of input channel attenuation to turn off unused microphones. The activation of an input channel becomes audibly apparent if the level change from the off state to the on state is too great. Practical experience has shown that a 15 dB change from off to on is a good compromise. However, as the number of microphones in the system increases, more input channel attenuation may be required for system gain stability. Adjustment of input channel attenuation is available on most automatic microphone mixers. This adjustment can be on an input-by-input basis or for all inputs at once. The relationship between gain before feedback, input channel attenuation, and the number of microphones is calculated by the following equation:

$$\Delta G = 10\log\frac{N}{1+(N-1)10^{A/10}} \qquad (21\text{-}4)$$

where,

ΔG is the gain improvement in dB with only one microphone activated,

N is the total number of microphones,

A is the attenuation for all input channels in dB.

Fig. 21-16 shows the relationship in graphical form. Note the asymptotic maximum value of gain improvement with infinite attenuation—i.e., all but one channel turned off. Also note that input channel attenuation greater than 30 dB offers little improvement for systems with up to 256 microphones.

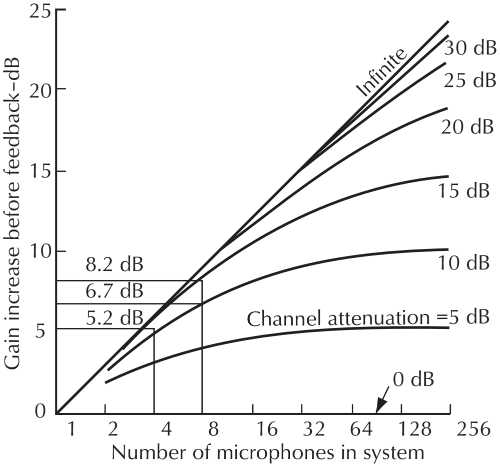


Figure 21-16. Gain improvement with different channel-off attenuations in a mixer that has a number of microphones and only one channel on.

21.3.3.7 Automatic Gain Control

Automatic gain control (AGC) of an input or output is a feature of a few automatic microphone mixers. A sound engineer rides gain to bring up weak signals or reduce overly loud signals and attempts to do this without destroying the inherent dynamic range of speech. An AGC in an automatic microphone mixer is typically designed to reduce gain only should the input signal level increase. The AGC is adjusted so that the quietest talker has maximum gain (without feedback). All louder talkers will force the AGC to bring down the overall level.

The IRP Level-Matic circuit is an example. It automatically adjusts the master gain to maintain a uniform output level for input signal variations up to 10 dB. A loud talker causes the gain to steadily decrease. When the talker stops, the gain holds as established by his or her average talking level. If a quiet talker then speaks, the gain steadily increases to a new value set by his or her average speaking level.

Gain control is based on loudness versus frequency and loudness versus time response of the ear. Gain adjustments are made at a constant dB per second rate to minimize the pumping and breathing effects of simple level compression circuits. If there is no signal, the AGC gain holds at its last value.

21.3.4 Types of Automatic Microphone Mixers

An automatic microphone mixer can have an analog circuit design, a digital circuit design, or a combination of the two. Though a digital design might offer more design flexibility due to software control, the digital automatic microphone mixer is not inherently better than an analog automatic microphone mixer. Be it analog or digital, an automatic microphone mixer will fall in one of the following functional groups:

1. Fixed threshold.
2. Variable threshold.
3. Gain sharing.
4. Direction sensitive.
5. Multivariable dependent.

21.3.4.1 Fixed Threshold Automatic Microphone Mixers

A detector circuit in the automatic microphone mixer activates an input channel when a microphone signal is present and attenuates the input when the microphone signal ceases. This basic function is often called a *noise gate*. To activate the input, the signal must be larger than a threshold preset for the channel during installation. This method has several shortcomings. First, there is the dilemma of where to set the activation threshold. If it is set too low, it will respond falsely to room noise, reverberation, and room-reflected sound. If the threshold is set too high in an effort to avoid false activation, desired speech signals may be chopped or clipped. The threshold should be set high enough to avoid activation by random noises, but low enough to turn on with desired speech signals. These are frequently contradictory requirements, and compromise is generally not satisfactory.

A more serious problem is that any number of input channels may activate with a very loud talker. One solution is a first-on inhibiting circuit that permits only one input channel to be on at a time. One-on-at-a-time operation is generally unacceptable for conversational dialog because the hold time needed to cover speech pauses will keep the second talker off.

Fixed threshold automatic microphone mixers have fallen out of favor and are now rarely employed. Early examples of fixed threshold activation products include the Shure M625 Voicegate (1973), the Rauland 3535 (1978), the Edcor AM400 (1982), and the Bogen AMM-4 (1985).

21.3.4.2 Variable Threshold Automatic Microphone Mixers

One attempt at overcoming the problems of a fixed threshold is to set the activation threshold based on a signal from a remote microphone. This microphone would be located in an area that is not expected to produce desired program input and is presumed to provide a reference signal that depends on variations in room noise or reverberation. Any desired talker input must then exceed this level by some preset amount. It is assumed that the desired talker signal will be louder than the reference. However, this may not be true, especially when the reference signal from a randomly selected microphone location does not represent the ambient sound in the vicinity of the talker's microphone. This is the basis of a system described by Dugan in U.S. Patent 3,814,856. An alternative source of reference threshold may be derived from the sum of the outputs of all the microphones in the system.

The discontinued JBL 7510 automatic microphone mixer employed a variable threshold design to override a fixed threshold. This design assumed that if a common acoustical disturbance was sensed at several microphone input channels, an input channel should not be activated. Instead, the overall system threshold should be raised. A talker must then be loud enough at the microphone to override the new raised threshold. Both the fixed threshold and the contribution of the background threshold reference would be set at installation. Release time, input attenuation, and gain were also necessary adjustments for each input channel. Variations on this concept of variable threshold design have been used in automatic microphone mixers from Audio Technica, Biamp, IED, Ivie, Lectrosonics, and TOA.

The Biamp autoTwo Automatic Mixer, Fig. 21-17, includes adaptive threshhold sensing to minimize false gate triggering, a speech frequency filter to minimize false gating due to noise, logic outputs from channels for switching external circuits, and 6 dB of hysteresis to reduce gate fluttering when near threshold. The block diagram is shown in Fig. 21-18.

Figure 21-17. Biamp autoTwo Automatic Mixer. Courtesy Biamp Systems.

21.3.4.3 Gain-Sharing Automatic Microphone Mixers

A gain-sharing automatic microphone mixer works from the premise that the sum of the signal inputs from all microphones in the system must be below some maximum value that avoids feedback oscillation. The safe system gain is set relative to the sum of all microphone signals in the system. If one microphone has more signal than the average of all signals, then that microphone channel is given more gain and all the other channels less gain roughly in proportion to the relative increase of signal level.

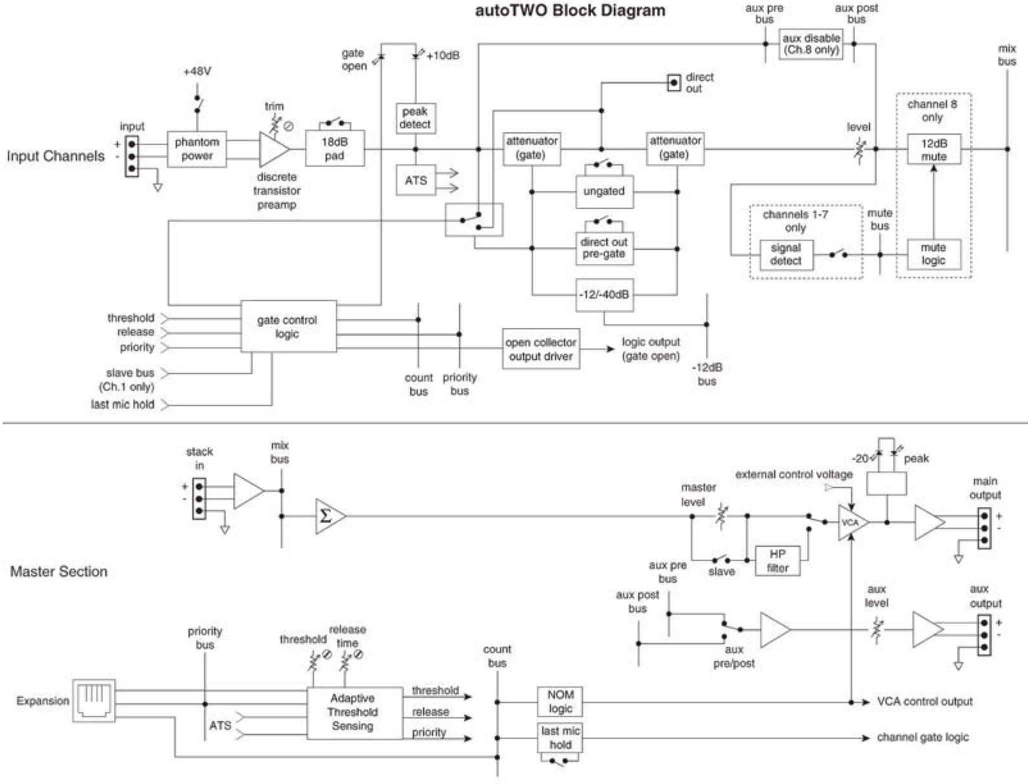

autoTWO Block Diagram
Input Channels
input
+
-
+48V
phantom power
trim
discrete transistor preamp
18dB pad
gate open
+10dB
peak detect
ATS
attenuator (gate)
ungated
direct out pre-gate
-12/-40dB
direct out
attenuator (gate)
aux pre bus
aux post bus
aux disable (Ch.8 only)
level
channel 8 only
12dB mute
mix bus
channels 1-7 only
signal detect
mute bus
mute logic
threshold
release
priority
slave bus (Ch.1 only)
last mic hold
gate control logic
count bus
priority bus
open collector output driver
logic output (gate open)
-12dB bus
Master Section
stack in
+
-
mix bus
Σ
master level
external control voltage
-20
peak
VCA
main output
+
-
aux pre bus
slave
HP filter
aux post bus
aux pre/post
aux level
aux output
+
-
priority bus
threshold
release time
Expansion
ATS
Adaptive Threshold Sensing
threshold
release
priority
count bus
NOM logic
last mic hold
VCA control output
channel gate logic

Figure 21-18. Block diagram of Biamp autoTwo Automatic Mixer. Courtesy Biamp Systems.

Dugan's U.S. Patent 3,992,584 describes such a system where a 3 dB level increase at one microphone causes that channel gain to go up by 3 dB, while the gain of the other channels decreases by 3 dB. Speech from two persons talking into separate microphones with levels differing by 3 dB (both appreciably above the background level) would appear at the output of the system with a 6 dB difference. In other words, the signal from a microphone with the highest output is given the most gain, and a signal from a microphone with the smallest output is given the least gain. With this operational concept, NOMA is not needed in the output stage. Theoretically, the system is configured so that the total gain is constant at a level that safely avoids feedback oscillation.

Automatic microphone mixers marketed by Dugan, Lectrosonics, Protech Audio, and Altec Lansing have used level proportional control based on average input signal amplitudes, Fig. 21-19.

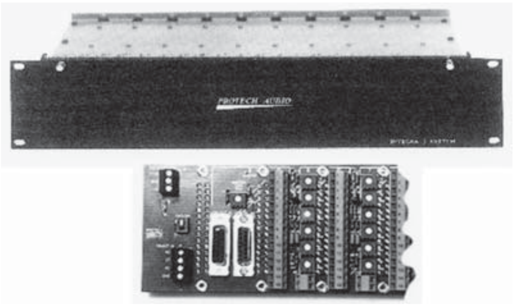

PROTECH AUDIO

Figure 21-19. Protech Audio automatic mixer. Courtesy Dan Dugan.

21.3.4.4 Direction-Dependent Automatic Microphone Mixers

A direction-dependent automatic microphone mixer responds to signals having acceptable levels within a predefined physical space in front of a microphone. By making the decision as to whether a channel should be on depends on the relative signal levels at two back-to-back cardioid microphone capsules in a single microphone housing. The Shure AMS (automatic microphone system) responds in part to the location of the sound source. This mixer works only with its own unique two-capsule microphones.

When an AMS input channel is activated, the front facing microphone signal is transmitted to the mixer output. This mixer functions like a variable-threshold system with its threshold being a fixed level above the background ambient noise but with the threshold also being a function of the sound source location and its angular relationship to the microphone.

Any input channel may turn on when the signal level from the front microphone capsule is 9.5 dB above the level from the rear capsule. Effectively, n SNE of 5 dB to 7 dB is required for a channel to activate. Of course, a weaker sound source will not activate the channel. The level difference of 9.5 dB is derived from the criterion that a cardioid microphone response at 60° off-axis is typically one-third of its on-axis response. The activation angle of the mixer input channel is thus 120°. A sound source outside of the 120° angle will not activate an input channel no matter what the sound pressure level.

To keep the AMS microphones compatible with conventional shielded twisted pair cable while keeping the two microphone signals separated, an unbalanced signal path is used. This approach can be more susceptible to induced hum and noise pickup than a conventional balanced signal path. The use of current source preamplifiers in the microphone and unusually low impedance inputs in the mixer minimizes this potential problem.

It is recommended that an AMS microphone be installed within three feet of each talker, and the talker must be located within the 120° activation angle. Each AMS microphone should also be at least three feet from any wall behind it and at least one foot from objects behind it such as books, large ashtrays, or briefcases. This precaution is necessary to avoid unwanted reflection of the talker's acoustic signal into the rear facing microphone capsule. Stray acoustic reflections can lead to unreliable input activation.

As the direction-dependent automatic microphone mixer process is covered under U.S. Patent 4,489,442, this type of automatic microphone mixer has been marketed only by Shure. In 2000, U.S. Patent 6,137,887 was issued to Shure for a new AMS design. Developed by Anderson, this patent adds a circuit that guarantees a single talker will activate only a single input channel, even if that talker is within the activation angles of multiple AMS microphones, Fig. 21-20.

Figure 21-20. Shure AMS8100 mixer. Courtesy Shure Incorporated.

21.3.4.5 Noise-Adaptive Threshold Automatic Microphone Mixers

This concept employs a dynamic threshold unique for each input channel. Using an inverse peak detector, each input channel sets its own minimal threshold that continually changes over several seconds based on variations in the microphone input signal.

Sound that is constant in frequency and amplitude, like a ventilation fan, will not activate an input but will add to the noise-adaptive threshold. Sound that is rapidly changing in frequency and amplitude, like speech, will activate an input. The mixer activates an input when two criteria are met:

1. The instantaneous input signal level from the talker is greater that the channel's noise-adaptive threshold.
2. The input channel has the maximum signal level for that talker.

Without this second criterion, a very loud talker might activate more than one input channel.

Note that this system deems any sound that is relatively constant in frequency and amplitude as nonspeech. Sustained musical notes may activate an input on attack, but after several seconds the sustained note will raise the threshold and the input will be attenuated. As previously stated, automatic microphone mixers are designed primarily for speech applications, not music.

Developed by Julstrom and covered by the U.S. Patent 4,658,425, the noise-adaptative threshold configuration has been used in automatic microphone mixers

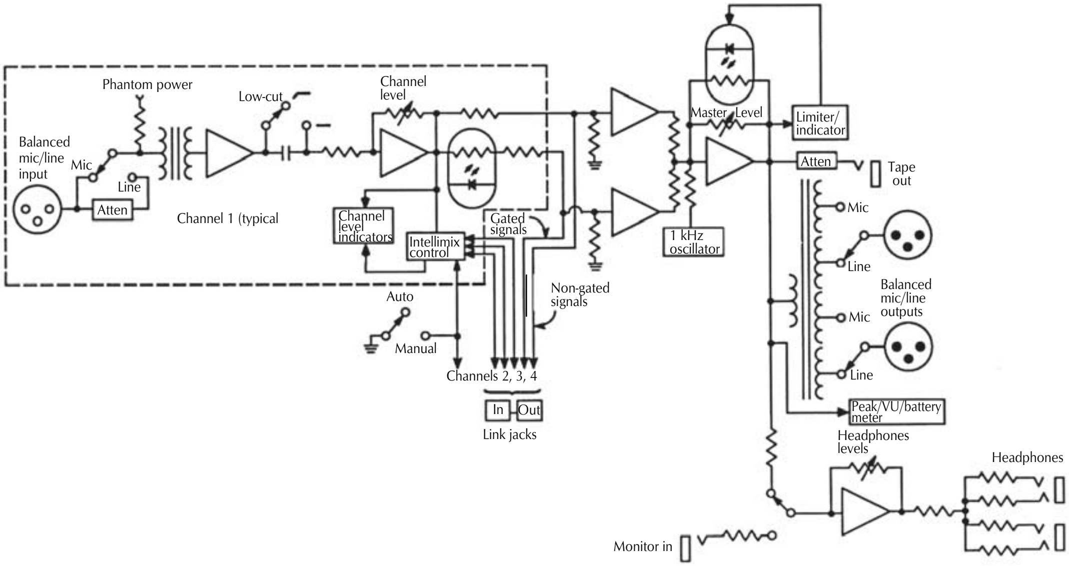

Phantom power
Low-cut
Channel level
Balanced mic/line input
Mic
Line
Atten
Channel 1 (typical
Channel level indicators
Intellimix control
Gated signals
Auto
Manual
Non-gated signals
Channels 2, 3, 4
In
Out
Link jacks
Master Level
Limiter/ indicator
Atten
Tape out
1 kHz oscillator
Mic
Line
Balanced mic/line outputs
Mic
Line
Peak/VU/battery meter
Headphones levels
Headphones
Monitor in

Figure 21-21. Block diagram of Shure FP410 mixer. Courtesy Shure, Incorporated.

manufactured by Shure, including the now obsolete FP410 battery-operated portable mixer, Fig. 21-21.

21.3.4.6 Multivariable-Dependent Automatic Microphone Mixers

The automatic microphone mixer methods described so far essentially use input signal amplitude as the activation variable. The relative timing of signals at each input is another variable that can be employed. A multivariable-dependent system makes its activation decision from both input signal amplitude and the time sequence of the input signals.

Peter's U.S. Patent 4,149,032 is such a design. The instantaneous positive signal amplitudes of all inputs are simultaneously compared to a threshold voltage (dc ramp) that falls 80 dB in 10 ms (or less) from a high value to a low value. Initially, all input channels are held in an attenuated state. The first input channel that has an instantaneous amplitude equal to the instantaneous value of the falling threshold is activated, while the other inputs remain attenuated. This activated channel remains so for 200 ms.

Once an input channel is activated, the threshold voltage is reset to its high value and immediately starts to fall again in search of another input to activate. If all talkers are silent and an amplitude match is not found, the threshold search progresses the full 80 dB in 10 ms and then resets. However, this scenario is not typical. Most of the time, a signal on one of the inputs will produce a threshold amplitude match early in the search. In practice, the average input activation time is 3 or 4 ms. Since the threshold resets every time an input is activated, the frequency of the threshold searches will be also every 3 or 4 ms on average.

As mentioned, the input activation is maintained for 200 ms. If on the second search the same input still has the largest signal amplitude, its activation status is renewed for another 200 ms. If during a future threshold search, a different input channel has the higher amplitude, it is activated for 200 ms. The first input activated times out and attenuates if not reactivated by a future search within the 200 ms. As long as a talker keeps speaking, his input is continually renewed for 200 ms intervals. This rapid response enables conversational dialog to be conducted and also permits easy activation of weaker sound sources during gaps in speech.

Since the activation gain of all input channels is the same, any signal source on an active channel has the same gain, and the relative levels of different talkers is preserved in the mixer output.

When multiple talkers vie for access to the system, the probability of all of them obtaining access decreases in proportion to the number. This effectively limits the maximum number of input channels that can be activated at any given time. For example, ten equally loud talkers will each be on 88% of the time. But as more than three or four persons talking at the same time is not intelligible, this limitation is normally of little consequence.

Also unique to the Peters design is the variable known as the *access ratio*. Simply put, the access ratio is the time an input is kept activated (200 ms) compared to the decision time taken to activate an input (10 ms). Access ratio may be readjusted to control the number of input channels that can activate at one time. Selective adjustment of the access ratio can also reduce missed beginnings of words.

21.3.4.7 Automatic Microphone Mixers with Matrix Mixing to Multiple Outputs

Recent designs in automatic microphone mixers have introduced matrix mixing to multiple outputs. This feature allows any input channel to be sent to any number of output channels and to be sent at different levels depending on the signal mix desired at the individual output. The Lectrosonics AM16/12 is a marriage of analog with digital. All control is accomplished via proprietary software that operates on a Windows-based computer. Software control allows a 16 in/12 out automatic microphone mixer with matrix mixing to fit in a two-rack space chassis. The software control also deters unauthorized readjustment as there are no knobs to twiddle.

Matrix mixing may be used for creating unique audio feeds for recording, teleconferencing, hearing assistance, language translation, etc. A courtroom is an example of a facility where all of these different audio systems might be required. Matrix mixing also provides the capability for mix-minus configurations. Simply put, a mix-minus output signal contains all input channels except for one or more—i.e., complete mix of all inputs minus one (or more) undesired inputs. The mix-minus concept improves gain before feedback. If a microphone signal does not appear in the closest loudspeaker, gain before feedback is better than if the microphone signal does appear in that loudspeaker. In a typical meeting room, talkers do not need to hear their own voice in the closest loudspeaker. They need to hear their other talkers located far away from their location. Mix-minus provides this capability, Fig 21-22.

However, matrix mixing also creates a problem—how to adjust so many gain variables? Consider an automatic microphone mixer with twelve inputs and one output. This mixer requires thirteen gain controls—one for each input and one for the master output. Now consider an automatic microphone mixer with 12 inputs and 8 outputs. This mixer requires 104 gain controls. One hundred four potentiometers and knobs take up a lot of panel space and are quite expensive. The answer to this problem is control via software. One example of this design concept is the Lectrosonics AM16/12.

21.3.4.8 Automatic Mixing Controller

The Model E-1 Automatic Mixing Controller, Fig. 21-23, helps professional audio mixers handle multiple live mics without having to continually ride their individual faders. This eight-channel signal processor patches into the input insert points of an audio mixing console. It detects which mics are being used and makes fast, transparent cross-fades, freeing the mixer to focus on balance and sound quality instead of being chained to the faders. The Model E-1's voice-controlled cross-fades track unscripted dialogue perfectly, eliminating cueing mistakes and late fade-ups while avoiding the choppy and distracting effects common to noise gates. Without the need for gating, a natural low-level room ambience is maintained.

Dugan automatic mixing controllers are used with multiple live mics and unscripted dialogue including talk shows, game shows, conference sound reinforcement, houses of worship, dramatic dialogue, wireless microphones in theaters, and teleconferencing. The Dugan controllers are typically connected in the insert points of the console's mic inputs, Fig. 21-24. Fig. 21-25 is the block diagram of the E-1 automatic mixing controller. Each unit handles up to eight channels, and the units can be linked together to accommodate a maximum of 64 mic channels.

The Model E-1 is an eight-channel line-level or ADAT digital insert device in a half-rack, one unit high cabinet and has minimal controls. Additional controls are available via a virtual control panel provided by an embedded web server. I/O is connected by TRS insert cables or ADAT optical cables. The Model E can be linked for up to 64 channels, and it can link with the Dugan Models D-2 and D-3. Power is 9–24 Vdc or 9–18 Vac.

Three models are available. The Model D-2 has analog I/O for use in the insert points of analog mixing consoles. The Model D-3 has AES digital connections for insertion into digital mixers. Both models feature a separate control panel that can be placed on the meter bridge or in front of the console.

21.3.4.9 Automatic Microphone Mixers Implemented in Software

If software can control automatic microphone mixer hardware, then automatic microphone mixers can also be completely created in software. This completely digital approach to automatic microphone mixers can be found in software based products offered by Allen & Heath, ASPI, BSS, Crown, Dan Dugan, Gentner, Lectrosonics,

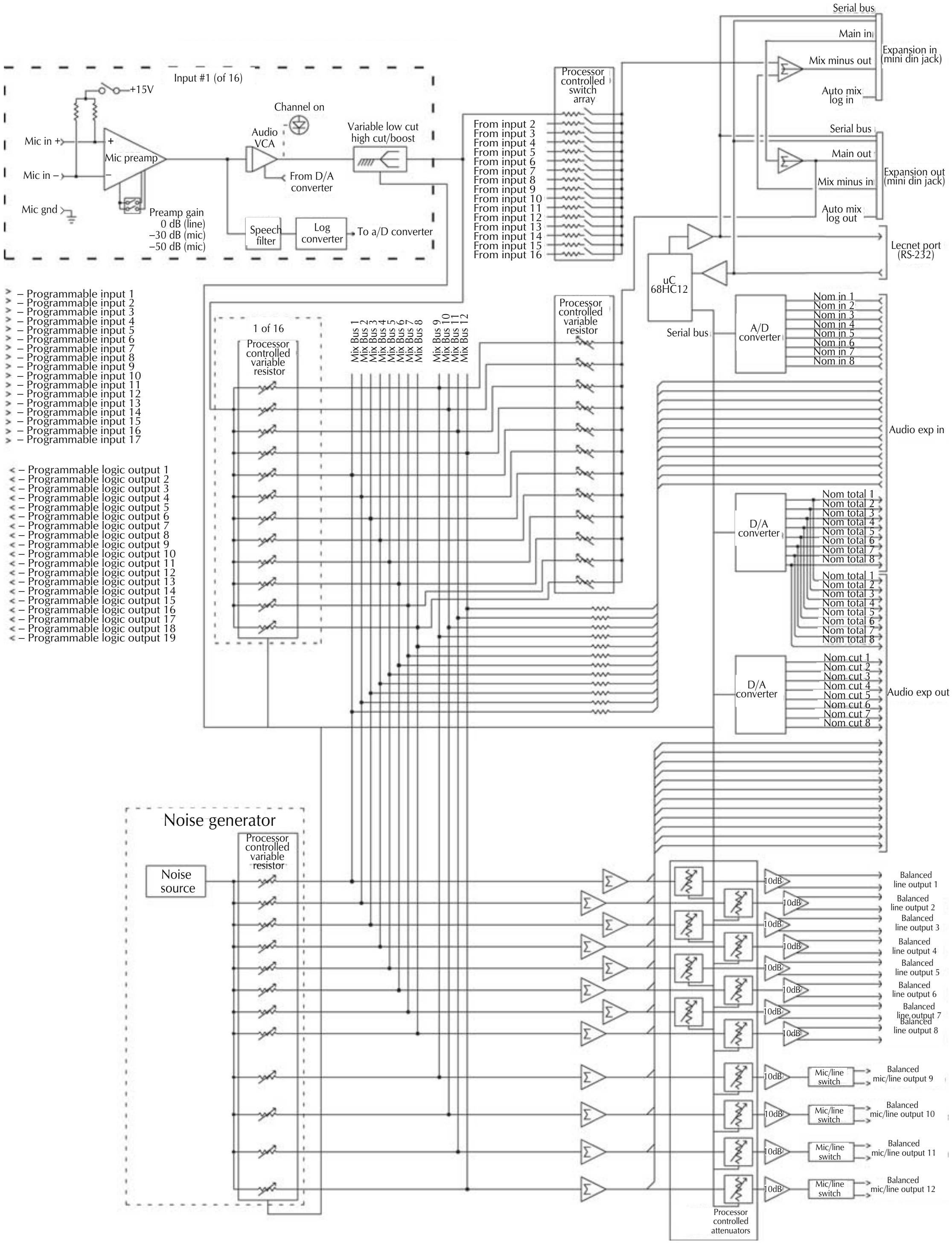
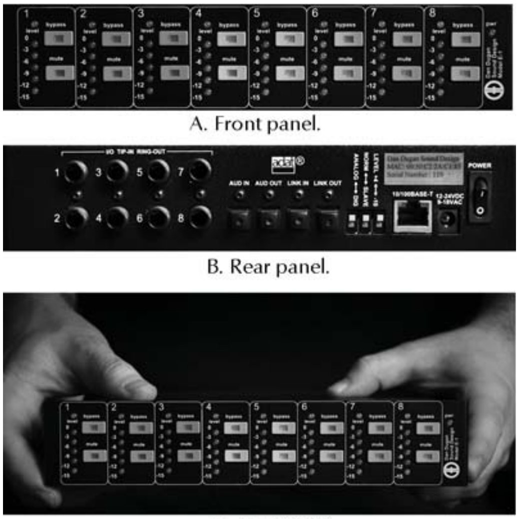

Input #1 (of 16)
+15V
Mic in +
Mic in –
Mic gnd
Mic preamp
Preamp gain
0 dB (line)
–30 dB (mic)
–50 dB (mic)
Channel on
Audio VCA
From D/A converter
Variable low cut high cut/boost
Speech filter
Log converter
To a/D converter
Processor controlled switch array
From input 2
From input 3
From input 4
From input 5
From input 6
From input 7
From input 8
From input 9
From input 10
From input 11
From input 12
From input 13
From input 14
From input 15
From input 16
Serial bus
Main in
Mix minus out
Auto mix log in
Expansion in (mini din jack)
Serial bus
Main out
Mix minus in
Auto mix log out
Expansion out (mini din jack)
Lecnet port (RS-232)
uC 68HC12
Serial bus
A/D converter
Nom in 1
Nom in 2
Nom in 3
Nom in 4
Nom in 5
Nom in 6
Nom in 7
Nom in 8
Audio exp in
Processor controlled variable resistor
> – Programmable input 1
> – Programmable input 2
> – Programmable input 3
> – Programmable input 4
> – Programmable input 5
> – Programmable input 6
> – Programmable input 7
> – Programmable input 8
> – Programmable input 9
> – Programmable input 10
> – Programmable input 11
> – Programmable input 12
> – Programmable input 13
> – Programmable input 14
> – Programmable input 15
> – Programmable input 16
> – Programmable input 17
< – Programmable logic output 1
< – Programmable logic output 2
< – Programmable logic output 3
< – Programmable logic output 4
< – Programmable logic output 5
< – Programmable logic output 6
< – Programmable logic output 7
< – Programmable logic output 8
< – Programmable logic output 9
< – Programmable logic output 10
< – Programmable logic output 11
< – Programmable logic output 12
< – Programmable logic output 13
< – Programmable logic output 14
< – Programmable logic output 15
< – Programmable logic output 16
< – Programmable logic output 17
< – Programmable logic output 18
< – Programmable logic output 19
1 of 16
Processor controlled variable resistor
Mix Bus 1
Mix Bus 2
Mix Bus 3
Mix Bus 4
Mix Bus 5
Mix Bus 6
Mix Bus 7
Mix Bus 8
Mix Bus 9
Mix Bus 10
Mix Bus 11
Mix Bus 12
D/A converter
Nom total 1
Nom total 2
Nom total 3
Nom total 4
Nom total 5
Nom total 6
Nom total 7
Nom total 8
Nom total 1
Nom total 2
Nom total 3
Nom total 4
Nom total 5
Nom total 6
Nom total 7
Nom total 8
D/A converter
Nom cut 1
Nom cut 2
Nom cut 3
Nom cut 4
Nom cut 5
Nom cut 6
Nom cut 7
Nom cut 8
Audio exp out
Noise generator
Noise source
Processor controlled variable resistor
10dB
Balanced line output 1
Balanced line output 2
Balanced line output 3
Balanced line output 4
Balanced line output 5
Balanced line output 6
Balanced line output 7
Balanced line output 8
Mic/line switch
Balanced mic/line output 9
Balanced mic/line output 10
Balanced mic/line output 11
Balanced mic/line output 12
Processor controlled attenuators

Figure 21-22. Block diagram of the Lectronics AM1612 mixer.



A. Front panel.

B. Rear panel.

C. E1-PS0202

Figure 21-23. Dugan E-1 Automatic Mixing Controller. Courtesy Dan Dugan Sound Design.

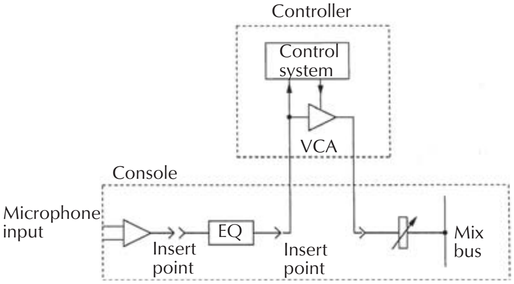


Figure 21-24. The automatic mixing controller is connected to the insert points of console input channels. Courtesy Dan Dugan Sound Design.

Peavey, and Rane. To date, the operational concepts used in digital automatic microphone mixers have not varied far from the previously described concepts underlying the analog automatic microphone mixers. This is likely to change, but as future digital automatic mixing concepts will be hidden deep within computer code, the manufacturers may be unwilling to reveal the details of operational breakthroughs; they will likely be kept as closely guarded company secrets. New concepts in automatic mixing might only become public knowledge if patents are granted or technical papers are presented.

The Polycom Vortex EF2280, Fig. 21-26, automatically mixes microphones and other audio sources while canceling acoustic echoes and annoying background noise. It is used in boardrooms, courtrooms, distance learning, sound reinforcement, and room combining. It connects easily to other equipment including codecs, VCRs, or other A/V products. The unit can be programmed from the front panel, or through Conference Composer™ software (included). Conference Composer's Designer™ wizard ensures fast, accurate setup for a variety of applications.

A single Vortex EF2280 unit provides automatic mixing of up to eight microphones plus four auxiliary audio sources. Up to seven additional Vortex EF2280 or Vortex EF2241 units can be linked to the first unit. (NOM) information can be specified across all channels in the linked units. The microphone channels feature acoustic echo cancellation to prevent retransmission of signals to their original locations. A neural network AGC reacts only to valid speech patterns, bringing voices within desired levels. AGC controls are user adjustable, as are settings for the five-band parametric EQ offered on all input and output channels and output delay controls. Fig. 21-27 is the block diagram of the Vortex EF2280.

21.3.4.10 Which Type of Automatic Mixer Works Best?

There is no definitive answer to this question. It is impossible to tell which automatic microphone mixer design will operate best in a given situation by studying technical specifications, believing the marketing literature, poring over circuit schematics, deciphering lines of computer code, or rereading this chapter. Human speech is very complex and human hearing is very discerning. Like so many areas in professional audio, the critical ear is the final judge.

21.3.5 Teleconferencing and Automatic Microphone Mixers

Automatic microphone mixers are used in many teleconferencing systems. The design of such systems involves a number of complex issues that do not enter into the design of sound reinforcement systems. This section will discuss important design aspects of such installations.

As practiced in modern communication between separated groups of talkers, teleconferencing has two components—visual and aural. The visual is handled by television cameras, video monitors, and video projectors. The visual may be full motion in real time, slow scan, or single-frame presentation.

It is appropriate to identify the aural part of the teleconferencing system as the audio conferencing system. Considerable attention must be paid to a number of details for acceptable sound quality, intelligibility, and user comfort. Users of teleconferencing systems tend to

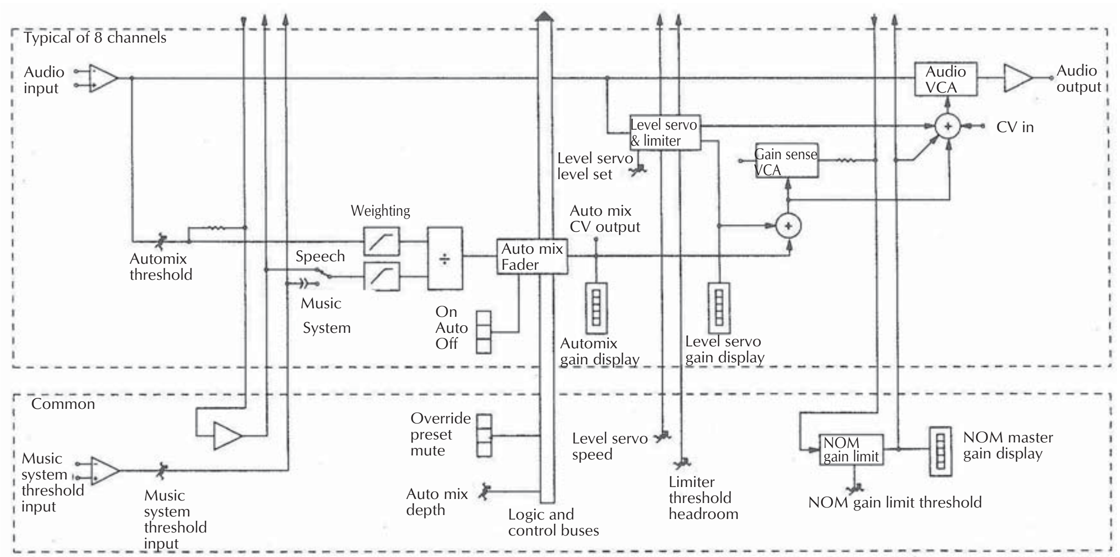

Typical of 8 channels
Audio input
Audio VCA
Audio output
Level servo & limiter
CV in
Level servo level set
Gain sense VCA
Weighting
Auto mix CV output
Automix threshold
Speech
Auto mix Fader
Music
System
On
Auto
Off
Automix gain display
Level servo gain display
Common
Override preset mute
Level servo speed
NOM gain limit
NOM master gain display
Music system threshold input
Music system threshold input
Auto mix depth
Logic and control buses
Limiter threshold headroom
NOM gain limit threshold

Figure 21-25. Block diagram of the E-1 automatic mixing controller. Courtesy Dan Dugan Sound Design.

EF2280
VORTEX
POLYCOM

Figure 21-26. Vortex EF2280 digital multichannel acoustic echo and noise canceller with a built-in automatic microphone/matrix mixer. Courtesy Polycom, Inc.

employ very subjective descriptions that have to be interpreted into quantitative engineering terms that can then be applied, measured, and included in system designs. Teleconference participants expect good speech intelligibility, easy identification of the talker, relatively high SNR, and other qualities. They also expect the overall aural experience to be better than a conversation conducted via telephone handsets.

An audio conferencing installation for voice and program has four primary facets:

1. Conference room and building acoustics.
2. Interface with telephone/transmission system.
3. Possible secondary use as a sound reinforcement system.
4. Proper equipment selection and setup.

21.3.6 Room and Building Acoustics

21.3.6.1 Conference Room Noise

The first consideration for a teleconference installation is noise in the room. Obvious noise sources, like heating and air conditioning systems, should be evaluated and specified for acceptable levels. External noise must also be considered:

- Conversations in hallways or adjacent offices.
- Business machines in adjacent spaces.
- Elevators on opposite sides of the wall.
- Water flow in building services.
- Vibration of air conditioners on the roof.***
- Loading docks

There will also be unwanted noise generated in the conferencing room itself:

- Fans in projectors and computers.
- Hum from light fixtures.
- Paper shuffling.
- Moving chairs.
- Coughing.
- Side conversations.

All of these undesired sound sources are much more obvious, annoying, and detrimental to intelligibility at the remote site of the teleconference than they are in the local site where they originate. Also, as the number of participants increases, the geographic area covered by the participants expands and unamplified speech becomes harder to hear due to greater distances between the talkers and the listeners. Consequently, for comfortable talking and listening, the ambient noise level in a room must be lower for larger groups.

Table 21-1 provides recommended noise level limits for conference rooms. These are levels at the conference

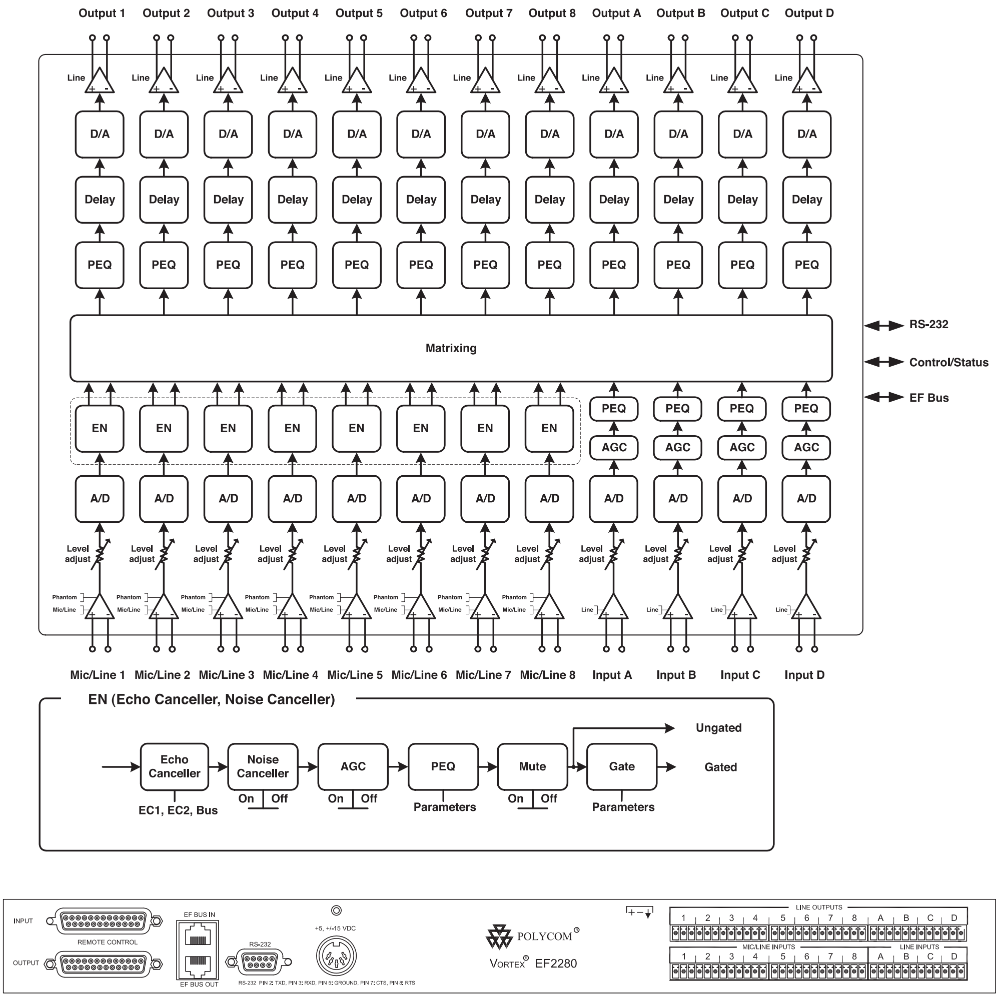

Output 1
Output 2
Output 3
Output 4
Output 5
Output 6
Output 7
Output 8
Output A
Output B
Output C
Output D
Line
D/A
Delay
PEQ
Matrixing
RS-232
Control/Status
EF Bus
EN
PEQ
AGC
A/D
Level adjust
Phantom
Mic/Line
Line
Mic/Line 1
Mic/Line 2
Mic/Line 3
Mic/Line 4
Mic/Line 5
Mic/Line 6
Mic/Line 7
Mic/Line 8
Input A
Input B
Input C
Input D
EN (Echo Canceller, Noise Canceller)
Echo Canceller
EC1, EC2, Bus
Noise Canceller
On
Off
AGC
On
Off
PEQ
Parameters
Mute
On
Off
Gate
Parameters
Ungated
Gated
INPUT
REMOTE CONTROL
OUTPUT
EF BUS IN
EF BUS OUT
RS-232
+5, +/-15 VDC
RS-232 PIN 2: TXD, PIN 3: RXD, PIN 5: GROUND, PIN 7: CTS, PIN 8: RTS
POLYCOM
VORTEX EF2280
LINE OUTPUTS
1 2 3 4 5 6 7 8 A B C D
MIC/LINE INPUTS
LINE INPUTS
1 2 3 4 5 6 7 8 A B C D

Figure 21-27. Block diagram of the Vortex EF2280 digital multichannel acoustic echo and noise canceller with a built-in automatic microphone/matrix mixer. Courtesy Polycom, Inc.

table with the room in normal unoccupied operation and at least 2 feet from any surface. Methods for achieving low interfering noise levels are discussed in Chapter 6, Small Room Acoustics.

More accurate assessment will result if noise criteria (NC) are used because of the strong influence of frequency spectrum on speech interference and listener annoyance (see Chapter 5, Acoustical Treatments for Indoor Areas).

Fig. 21-28 shows the maximum microphone/talker distance for a marginally acceptable SNR of 20 dB in transmitted speech. The graph applies to omnidirectional microphones. The distance may be increased by 50% for directional microphones. If more than one microphone is active in the system, the number of open microphones must be taken into account by reducing the predicted SNR by 3 dB for each time the number of open microphones doubles. An automatic microphone mixer will alleviate this concern.

If an acoustical survey indicates the presence of interfering noise sources, construction techniques must be implemented to provide adequate sound transmission losses, or another room should be considered.

21.3.6.2 Conference Room Reverberation

Reverberation is often identified by conference participants as the “speaking into a barrel” effect. The sources of reverberation are variable. For example, there is the reverberation from hard surfaces in the room where

Table 21-1. Ambient Noise Level Limits for Conference Rooms

Conference Size	Maximum Sound Level in dBA	Preferred NC	Acoustic Environment
50 people	35	20–30	Very quiet, suitable for large conferences at 20–30 ft table.
20 people	40	25–35	Quiet, satisfactory for conferences at a 15 ft table.
10 people	45	30–40	Satisfactory for conferences at 6–8 ft table.
6 people	50	35–45	Satisfactory for conferences at 4–5 ft table.

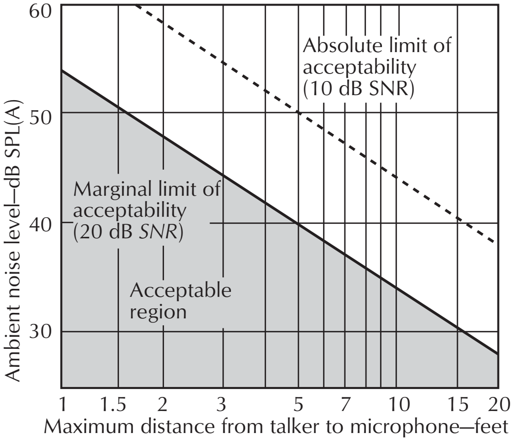


Figure 21-28. Acceptable ambient noise levels.

speech is originating at the moment. Requirements for comfortable listening in the room dictate reverberation times that are not too short, as a bit of acoustic liveliness in a meeting room is desirable. If an automatic microphone mixer is not employed to reduce the reverberation picked up by unused microphones, the reverberation heard at the remote site of the teleconference can be excessive and intolerable. In this situation, a very low (and uncomfortable) reverberation time at the local site is required. The potential for reduced intelligibility at the remote site is increased because the remote participants do not have the advantage of separating the speech signal from the reverberation via binaural hearing, plus not having the talker in the same room also tends to dull one's attention.

There is also the reverberation added at the remote site. The incoming signal is reproduced by loudspeakers, the sound propagates around the room, and even more reverberation is added to the talker's signal. So, unless there is only a telephone handset at the remote site, both sites need to have proper acoustical characteristics. This is often not the case as in many conferencing rooms the visual comfort of the room takes precedence over the aural comfort. Just ask the interior designer!

Room dimensions should be chosen to minimize standing waves and flutter echoes. If the room already exists, judicious use of acoustically absorbent material is advisable for control of the room's acoustics.

Critical distance (D_c) is often used to predict appropriate talker to microphone distances. D_c is where the direct signal of the talker is equal to reflected signal of the talker, i.e., 50% direct signal and 50% reverberant signal. The D_c in conference rooms is typically in the range of 1–4 feet. D_c may be estimated from reverberation time measurements:

$$D_c = 0.03\frac{V}{T} \qquad (21\text{-}5)$$

where,

D_c is the critical distance in feet,
V is the room volume in ft³,
T is the reverberation time in seconds for 60 dB decay.

For good intelligibility, an omnidirectional microphone should be placed at ½ of D_c or less from the talker. When a directional microphone is used, the distance between talker and microphone may be increased up to 75% of the critical distance.

Because the sound decay in the first 60–100 ms usually is the most damaging to teleconference conversations, the usual reverberation time measurement, RT_{60}, may not be the most appropriate. One manufacturer of conference equipment insists that the room produce a decay of greater than 16 dB in the first 60 ms.

The one admonition to anyone faced with the design of a teleconferencing system is *do not ignore the acoustical characteristics of the room*. Insist upon a room that has the right acoustical environment or commit the resources to make it right before proceeding with the rest of the project. Acoustical deficiencies can rarely be corrected by electronic means. If it is new construction, work closely with the architect before the room design is complete.

21.3.6.3 Telephone/Transmission System Interface

Fig. 21-27 shows two teleconference rooms connected by a single two-wire telephone line. Each room has a

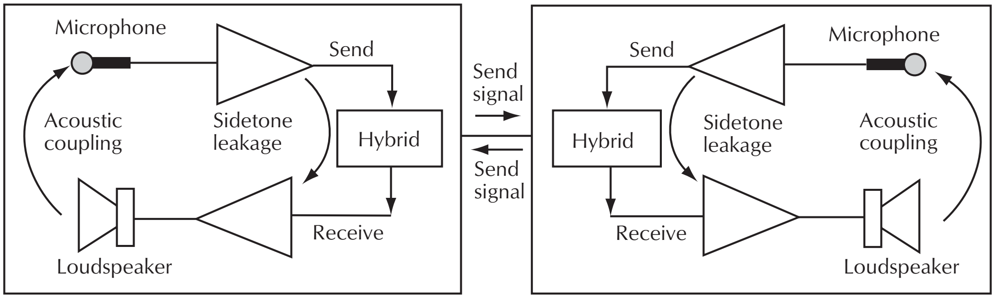


Figure 21-29. Teleconference system with two-wire telephone connection, showing feedback paths.

microphone and a loudspeaker with associated amplification. A *hybrid* interface between the send and receive lines and the telephone line serves to reduce loop gain within the room by reducing *sidetone leakage*.

Possible feedback loops are shown. Not only is there potential oscillation in the sending room, but also the coupling through the line to the receiving room and back is an equally probable feedback loop. Basic speakerphones use voice-activated gates to capture the line and permit transmission in only one direction at a time and thus interrupt the feedback path from the remote site. This can cause frequent dropouts in a conversation and forces the communication into a *half-duplex* mode of operation. Half duplex transmission refers to transmission in only one direction at a time.

The ideal is *full duplex*, which allows transmission in both directions all the time. A phone call from telephone handset to telephone handset provides full duplex communication. Full duplex is preferred for audio conferencing because there are no missing words or sentences, and conversations can be conducted in a normal manner. Control of reverberation and room noise is essential in any full duplex system.

An alternative connection system uses four wires as shown in Fig. 21-28. One pair of wires is used for each direction of transmission, thus eliminating the often troublesome hybrid sidetone leakage. As can be seen, there is still the possibility of feedback through the room at either end. However, there is usually cleaner signal transmission with the added expense of a second telephone line. Four-wire systems make full duplex communication possible.

Frequently audio conferences involve several sites giving rise to point-to-multipoint or multipoint-to-multipoint telephone interconnections. A *conference bridge* is used to connect a number of telephone lines so that all participants will be tied together. Bridging over 20 phone lines is now quite common. The actual bridging may be provided by an external bridging service company or bridging devices may be part of the on-site teleconferencing equipment.

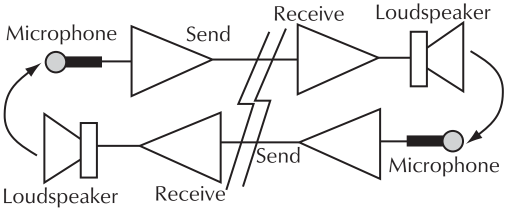


Figure 21-30. Teleconference system with a four-wire telephone connection.

A typical conference bridge limits the number of open ports to two because signal leakage in the bridge can cause retransmission of received audio on telephone lines. As a result, only one two-way conversation can occur and others can only listen. Also, the uncertain and variable quality of telephone connections can result in having a noisy line tying up the system and preventing access since the bridging control depends on signal-activated switching.

21.3.7 Teleconferencing Equipment

21.3.7.1 Telephone Interface

The telephone interface for a typical two-wire site is the hybrid. It converts the two-wire transmission of the connecting lines to internal four-wire paths to isolate the send and receive signals. A hybrid passes the microphone send signal (two of the four wires within the room) to the two-wire telephone line but attenuates it to the receive line. Conversely, a signal being received from the telephone line passes to the receive line (the other two of the four wires in the room) and is attenuated to the microphone send line. For many years, the hybrid in a standard telephone set was a transformer; now electronic equivalents are common.

The conference bridge operates in a similar manner. Good balance in both the bridge and the hybrid is necessary. This involves well-controlled and constant telephone impedance. Unless these devices can adapt to variable telephone line conditions, signal leakage may be retransmitted through them. If the boardroom has not been correctly treated acoustically, the combination of room echo and signal leakage creates an undesired feedback path. Many hybrids suppress leakage by less than 15 dB whereas 35 dB to 40 dB is regarded as the minimum acceptable for loudspeaker receive conference installations. The paths of signal leakage are shown in Fig. 21-31. Active hybrids are supplied by manufacturers such as Gentner Electronics, ASPI, and Telos. These products provide means for optimizing the impedance match to the telephone line, thereby giving additional suppression of signal leakage. Active hybrids can make the difference between a marginal and an acceptable teleconference.

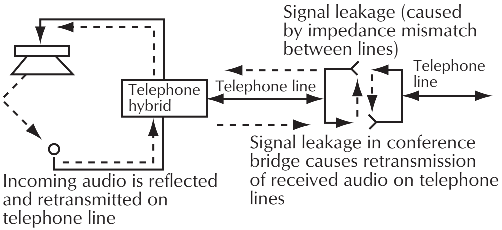

Signal leakage (caused by impedance mismatch between lines)
Telephone line
Telephone hybrid
Telephone line
Signal leakage in conference bridge causes retransmission of received audio on telephone lines
Incoming audio is reflected and retransmitted on telephone line

Figure 21-31. The paths for signal leakage and undesired feedback in a typical teleconference system.

Typical telephone line impedances range from 600 Ω to 900 Ω. Telephone equipment expects send levels of 0 dBm. The receive level standard is –6 dBm, but these levels are reported to vary widely, –10 dBm is frequently experienced. The standard telephone line has 48 Vdc (some private exchanges use 24 Vdc) for system control that must be blocked with a transformer or capacitors. The dc current through the off-hook relay keeps the line open while the connection is active.

21.3.7.2 Microphone Considerations

For a small group in a conference room, it may be possible to use only one omnidirectional microphone on a table top, typically of the surface-mount type. However, even for a group of four to six people, the equivalent of several directional microphones with an automatic microphone mixer is preferred to reduce the number of open microphones to the minimum necessary for the discussion. Three cardioid microphones in a circle, spaced at 120° intervals, is a typical approach. There are a number of surface-mount microphones that can be used, provided that the distance to the talkers is acceptably short. The typical participant in a teleconference expects, at minimum, the sound quality heard from a handset where the microphone is within inches of the talker's mouth. Thus, keeping the microphones close to the talkers is very important.

21.3.7.3 Microphone Mixing

Larger groups inevitably require a large number of microphones to keep the participant-to-microphone distance within the limits set by the room's critical distance. Some form of automatic microphone selection and mixing is essential in this case. Systems can be designed using an automatic microphone mixer (as described earlier in this chapter) connected to a telephone line interface device. Or a system can be implemented using an integrated device where the automatic mixing and the telephone interface are contained in the same chassis. Depending on the complexity required, there are many suitable approaches to system design. Consult the equipment manufacturers for specific design suggestions, Fig. 21- 32.

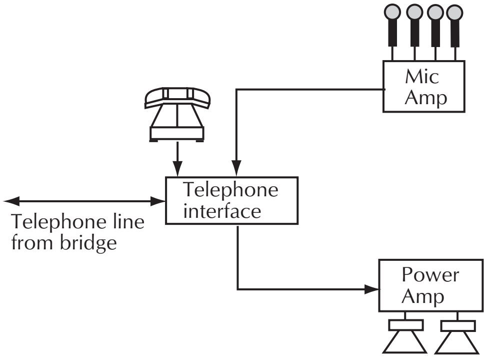


Figure 21-32. The configuration of a multimicrophone audio conference installation, without sound reinforcement.

21.3.7.4 Loudspeaker Considerations

Direct feedback from loudspeakers to microphones in any of the conference sites must be avoided; therefore, loudspeaker placement is critical. Loudspeakers should be placed in the null of the microphone pickup patterns. For cardioid microphones pointing in a horizontal direction, loudspeakers can be placed behind the microphones and aimed upward. Never place loudspeakers in front of microphones as microphones cannot distinguish between talkers in the room (desired sound sources) and talkers heard via loudspeakers (undesired sound sources).

When there is talking in the room, automatic microphone mixers can reduce the level of the loudspeaker signal from the remote site. This is accomplished via attenuating relays, ducking circuits, etc. By contrast, Sound Control Technologies offers a system that places the loudspeaker symmetrically between a pair of microphones that are out of polarity with each other. The loudspeaker contribution to the send line is claimed to be reduced by 40 dB with this arrangement.

If sound reinforcement of conversations within the room (sometimes known as voice lift) must also be provided in addition to audio conferencing, even more attention must be given to reducing the audio coupling between the loudspeakers and the microphones. Such systems can be very difficult to design correctly and must be approached with great caution. The use of an experienced acoustical/audio consultant is highly recommended in these cases.

21.3.7.5 Send Level Control

Send level—i.e., the audio signal voltage supplied to the telephone line—should be within acceptable ranges. Compressors, AGCs and levelers are all devices to consider for this technical requirement.

21.3.7.6 Echo Canceler

Echo cancelers reduce residual echo return in audio conferencing installations. If the local site returns significant signal from its incoming port to its outgoing port, and there is significant propagation delay due to the transmission line, the remote site will hear an annoying echo when someone in the remote site speaks.

The imperfect balancing of hybrids is one path for echo. Signal reflection within the telephone line is another source of echo. Echo also occurs acoustically when loudspeaker sound reaches open (active) microphones that are transmitting speech. The use of satellite transmission links also makes echo problems worse because of the long propagation delays.

A line echo canceler attempts to reduce echoes that are electronic in nature, such as those caused by hybrid leakage. An acoustic echo canceler looks at the signal coming into a room and inserts a time-delayed mirror image of the incoming signal into the outgoing signal leaving the room. The idea is to cancel any of the incoming signal that leaks into the outgoing signal path as a result of the acoustical coupling between loudspeaker and microphone.

Echo canceler technology has rapidly advanced due to faster CPU speeds and new research into canceler algorithms. Early echo cancelers were very expensive and thus having a single canceler at each conferencing site was considered adequate. As the price of echo cancelers has declined, manufacturers such a Gentner and ASPI now offer devices that have an echo canceler for each microphone input channel.

21.3.7.7 Historical Examples of Teleconferencing Equipment

Two historical systems will be described in more detail in order to show the number of parameters that must be considered in addition to the usual sound reinforcement needs. The first is an automatic microphone mixer approach as exemplified by the Shure ST3000, first manufactured in the 1980s. The second is the Sound Control Technologies system that does not use automatic microphone mixing.

Shure ST3000—An Analog Speakerphone. A simplified block diagram of the ST3000 is shown in Fig. 21-33. A conference call connection is made by taking the telephone handset from its cradle and dialing the desired number. When it is determined there is a good connection with the dialed party, the controller conference switch can be depressed to turn on the conference system. Green talk LEDs turn on and the handset may be returned to its cradle. The controller loudspeaker volume may next be adjusted if necessary. Levels for any auxiliary equipment may also be adjusted. Use the mute switches to prevent the called party from hearing local conversation. Red LEDs indicate muted status. The conference is terminated by depressing the controller telephone switch for at least 1 second.

In Fig. 21-33, the upper left mixer amplifier feeds the various auxiliary outputs. Below this amplifier, the conference microphone inputs are shown. Only the mutable (i.e., for automatic mixing) microphone inputs feed the send signal path to the hybrid. The receive path leads to the power amplifier and loudspeaker. Relative send and receive signals in the room are controlled by the send/receive switching and suppression logic. The suppression logic causes either the send amplifier or the receive amplifier to attenuate its signal depending on the presence of a receive signal. Because standard voice quality telephone lines have restricted bandwidth requirements, bandpass filtering is included in the send

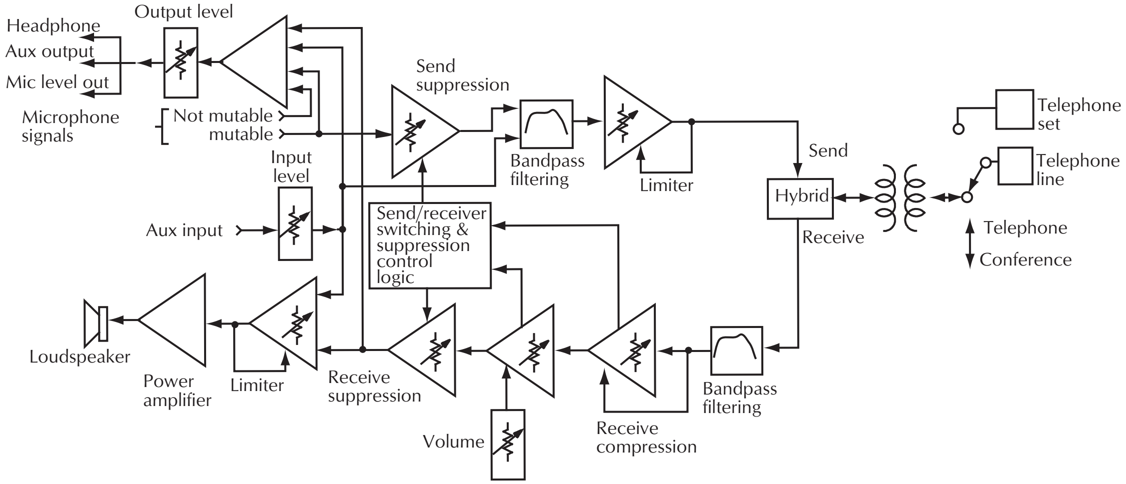

Headphone
Aux output
Mic level out
Output level
Microphone signals
Not mutable
mutable
Input level
Aux input
Send suppression
Bandpass filtering
Limiter
Send
Hybrid
Receive
Telephone set
Telephone line
Telephone
Conference
Send/receiver switching & suppression control logic
Loudspeaker
Power amplifier
Limiter
Receive suppression
Volume
Receive compression
Bandpass filtering

Figure 21-33. Block diagram of an analog teleconferencing system using automatic microphone control techniques. Courtesy Shure Incorporated.

channel. Bandpass filtering in the receive channel reduces the possibility of extraneous noise from the telephone line.

In the early 1990s, digital technology replaced analog devices such as the Shure ST3000. Polycom is now one of the prime suppliers of digital, full duplex speakerphones.

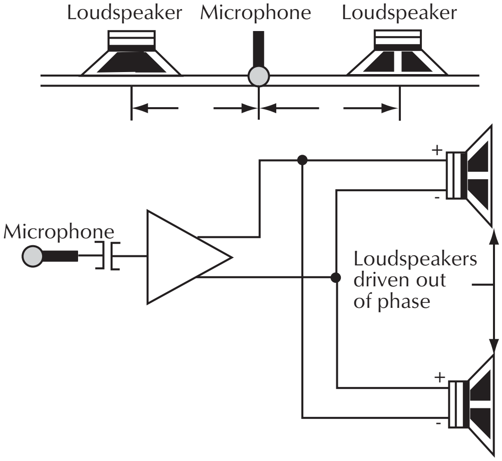

Loudspeaker
Microphone
Loudspeaker
+
-
Microphone
Loudspeakers driven out of phase
+
-

Figure 21-34. A method of using acoustic cancellation to reduce acoustic leakage in a conference room by driving loudspeakers out of phase.

Sound Control Technologies Ceiling Systems. Two configurations have been supplied by Sound Control Technologies. Loudspeakers and microphones are mounted in the ceiling over the conference participants. In one configuration, two loudspeakers are driven in antiphase (180% out of polarity) and a small microphone is mounted midway between them. Direct sound from the loudspeaker to the microphone is balanced for a null of 20 dB for receive signals. The basic element is shown in Fig. 21-34.

The second configuration uses a microphone and a loudspeaker mounted precisely 12 inches apart in reflecting baffle ceiling-mounted panels. Pairs of these loudspeaker/microphone units are placed above the conference table. All loudspeakers are driven in the same phase, while the microphones of symmetrically located units are mixed and balanced antiphase. A block diagram of such a system is shown in Fig. 21-35.

The microphone signals being mixed and balanced in antiphase feed the bus from which both the sound reinforcement (voice lift) and telephone send signals are derived. Notch filters are used for adjustment of spectrum balance. Delay may be included in the reinforced sound feeds if the room is large. The telephone return signal also feeds the sound reinforcement loudspeakers. An echo canceler is included to reduce the effects of telephone line echo or room acoustic echo.

As with the Shure system described previously, a telephone connection is made with a handset. Upon completion of the connection, the status of the line is determined by transmission of a group of tone bursts that allows the hybrid to electronically balance for the complex impedance of the telephone line. A push-button switch converts the connection to conference and the handset may be placed in its cradle.

21.3.7.8 The Present and Future of Teleconferencing

Most basic teleconference systems are sophisticated speakerphones with full duplex capability. Mid-level teleconferencing systems employ automatic microphones mixers and digital hybrids. The most sophisticated systems feature integrated teleconferencing devices that include multiple inputs with automatic mixing and echo cancellation, mix-minus signal routing capability, real-time feedback and level control, and operation via touchscreen.

Personal computers and digital signal processing (DSP) are becoming the dominant technologies that drive new developments in teleconferencing.

DSP advances are leading to teleconferencing systems that provide each participant with a customized electro-acoustical environment, unique to his or her own talking and hearing requirements. Advances in background noise reductions via electronic means are already impressive, as long as the noise has a repetitive nature. Microphone arrays that can be steered to best pick up a talker and steerable loudspeaker arrays are more prevalent.

But no matter how dominant digital technology becomes in teleconferencing, the speech input to the system from the human mouth will be analog, and the acoustical output to the human ear will also be analog. And that is the only technology forecast that will be 100% accurate.

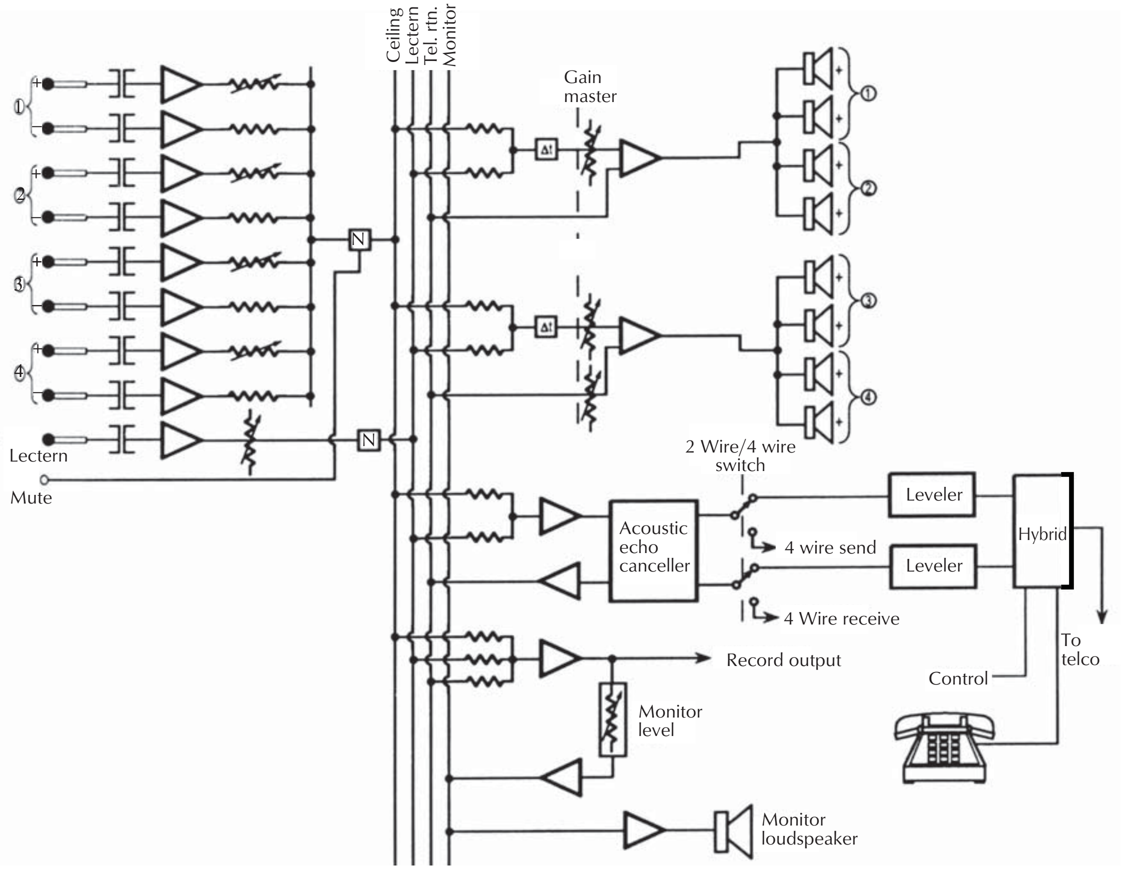

Ceiling
Lectern
Tel. rtn.
Monitor
Gain master
①
②
③
④
N
Δt
Lectern
Mute
2 Wire/4 wire switch
Leveler
Leveler
Hybrid
Acoustic echo canceller
4 wire send
4 Wire receive
Record output
Control
To telco
Monitor level
Monitor loudspeaker

Figure 21-35. Schematic of an teleconference system in a board room that uses all loudspeakers in phase and pairs of microphones in antiphase. Courtesy of Sound Control Technologies, Inc.

References

1. C. Motchenbacher and F. Fitchen, *Low-Noise Electronic Design*, New York: John Wiley & Sons, 1973, p. 10.
2. Ibid, pp. 34—35.
3. D. Jensen, *High Frequency Phase Specifications— Useful or Misleading?* AES 81st Convention Preprint 2398, 1986, pp. 5–9.
4. W. Leach, Jr., "The Differential Time-Delay Distortion and Differential Phase-Shift Distortion as Measures of Phase Linearity," *AES Journal,* Vol. 37, no. 9, 1989, pp. 712–714.
5. D. Jensen and G. Sokolich, *Spectral Contamination Measurement*, AES 85th Convention Preprint 2725, 1988.

Patent References

D. Dugan, U.S. Patent 3,814,856; June 1974.

D. Dugan, U.S. Patent 3,992,584; November 1976.

R. Peters, U.S. Patent 4,149,032; April 1979.

R. Ponto, U.S. Patent 4,239,938; December 1980.

R. Ponto, H. Martin, U.S. Patent 4,374,300; February 1983.

C. Anderson, W. Bevan, R. Schulein, A. Smith, U.S. Patent 4,489,442; December 1984.

S. Julstrom, U.S. Patent 4,658,425; April 1987.

S. Sakamoto, U.S. Patent 5,148,491; September 1992.

T. Sims, Jr., U.S. Patent 5,402,500.

T. Sims, Jr., U.S. Patent 5,414,776.

Anderson, M., U.S. Patent 6,137,887; October 2000.

Bibliography

"A Survey of Automatic Mixers," F. Ampel; *Sound & Video Contractor*; October 1983.

"Acoustics of Teleconference Rooms," Bell Labs, Publication 42901, August 1977.

"Acoustics of Teleconferencing," C. Stockbridge, *Journal of the Acoustical Society of America,* vol. 74, no. 81, p. 85.

"Automatic Microphone System Updates," M. Pettersen; Shure Inc.; 34 technical newsletters; 1983–1995.

"Automatic Mic Mixers," R. Baum; *Sound & Communications*; series in four parts: July 1993, September 1993, November 1993, March 1994.

"Broadcast Applications for Voice Activated Microphones," M. Pettersen; National Association of Broadcasters Conference Proceedings; May 1989.

"Conference Audio Systems Design," M. Pocino, *Sound & Video Contractor*, March 2001.

"Echo Cancellation Explained," S. Wooley, *Sound & Communications*, January 2000.

"Preamplifiers and Mixers," M. Burkhard; *Handbook for Sound Engineers*; 2nd edition; Howard W. Sams; 1991.

"Room Acoustics and Their Effect on Microphone Placement," S. Julstrom and M. Pettersen, Invited paper, Acoustical Society of America, 118th meeting, November 1989.

"The Theory Behind Automatic Mixers," D. Dugan; *Sound & Video Contractor,* March 1994.

This page intentionally left blank

Chapter 22

Attenuators

by Glen Ballou

22.1 General 767
  22.1.1 Attenuator K Value 768
  22.1.2 Loss 769
  22.1.3 Impedance Matching 769
  22.1.4 Installations, Practices, and Measurements 770
22.2 Types of Attenuators 771
  22.2.1 L Pads 771
  22.2.2 Dividing Networks 772
  22.2.3 T Attenuators 773
  22.2.4 Bridged T Attenuators 773
  22.2.5 π or Δ Attenuators 774
  22.2.6 U Attenuators 775
  22.2.7 Ladder Attenuators 776
  22.2.8 Simple Volume and Loudness Controls 776
  22.2.9 Light-Dependent Attenuators 778
  22.2.10 Feedback-Type Volume Control1 778
  22.2.11 Voltage-Controlled Amplifiers 779
  22.2.12 Field Effect Transistor Attenuators 779
  22.2.13 Automated Faders 779
  22.2.14 Automatic Attenuators 779
  22.2.15 Mixers 780
  22.2.16 Summing Amplifiers1 781
Reference 782

This page intentionally left blank

22.1 General

Most of the circuits today do not require passive attenuators and/or impedance matching devices as their input impedance is high and their output impedance is low. However, if a low-impedance output feeds a long line to a high-impedance input, high-frequency losses will occur if the line is not terminated with a matched impedance. This may be thousands of feet or a few feet when using older equipment that was designed for matched operation. When connecting to external circuits, the signal must often be attenuated to meet standards, a good place for low-maintenance passive attenuators.

An attenuator or pad is an arrangement of noninductive resistors in an electrical circuit used to reduce the level of an audio- or radio-frequency signal without introducing appreciable distortion. Attenuators may be fixed or variable and can be designed to reduce the signal logarithmically or any other curve.

Attenuator networks have been in use since the inception of the telephone for controlling sound levels and the matching of impedances. Many of the present-day configurations are the work of Otto J. Zobel, W. H. Bode, R. L. Diezold, Sallie Pero Mead, and T. E. Shay, all of the Bell Telephone Laboratories. Also, tables of constants developed by P. K. McElroy (also of Bell Telephone Laboratories) for various values of expression and substitution in equations have long been time-savers for the design engineer.

Attenuators and pads may be *unbalanced* or *balanced*. In an unbalanced attenuator, the resistive elements are on one side of the line only, Fig 22-1. In the balanced configuration, the resistive elements are located on both sides of the line, Fig. 22-2.

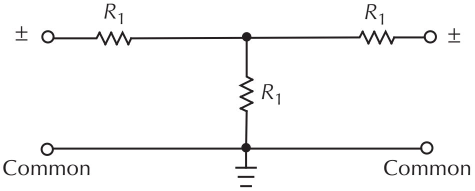


Figure 22-1. An unbalanced T-type attenuator

An *unbalanced pad* should be grounded to prevent leakage at the higher frequencies. The line without the resistor elements, called the *common*, is the only line that should be grounded. If the side with the resistors is grounded, the attenuator will not work properly, in fact, the signal will probably be shorted out.

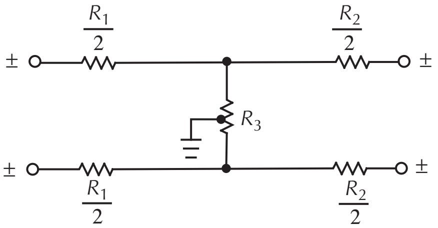


Figure 22-2. A balanced T-type attenuator.

A *balanced attenuator* should be grounded at a center point created by a balancing shunt resistance.

Balanced and unbalanced configurations cannot be directly connected together; however, they may be connected by the use of an isolation transformer, Fig. 22-3. If the networks are not separated electrically, half of the balanced circuit will be shorted to the ground, as indicated by the broken line in Fig. 22-4. Here severe instability and leakage at the high frequencies can result. The transformer will permit the transfer of the audio signal inductively while separating the grounds of the two networks. Even if the balanced network is not grounded, it should be isolated by a transformer. Transformers are usually designed for a 1:1 impedance ratio; however, they have taps for other impedance ratios. Chapter 32, Grounding, discusses the proper way to connect equipment to eliminate ground problems.

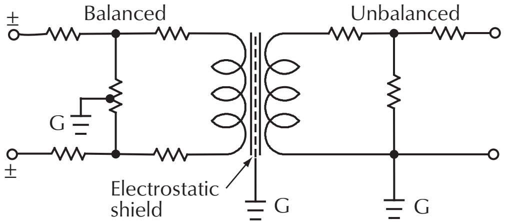


Figure 22-3. Correct method of connecting balanced and unbalanced networks through a transformer.

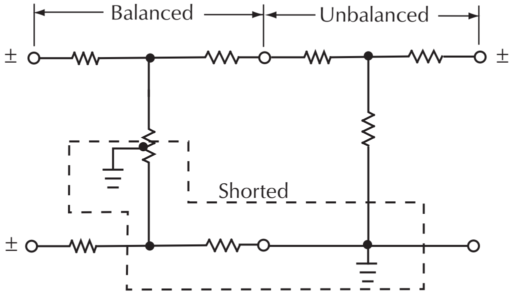


Figure 22-4. Two networks, one balanced and one unbalanced, connected incorrectly.

22.1.1 Attenuator K Value

To simplify the design of complex attenuators, a *K* value is used in the equation. *K* is the ratio of current, voltage, or power corresponding to a given value of attenuation expressed in decibels. The equation for *K* is

$$K = 10^{\frac{dB}{20}} \tag{22-1}$$

$$K = 10^{\frac{dB}{10}} \tag{22-2}$$

To simplify the calculation of attenuator networks, the values of the most frequently used expressions as tabulated by P. K. McElroy are given in Table 22-1. The various values of the expressions are substituted in the equations, saving much time.

Table 22-1. K″ Factors for Calculating Attenuator Loss Values

	a	**b**	**c**	**d**	**e**	**f**	**g**	**h**	**i**	**j**	**l**	
n (dB)	$r = \frac{1}{K}$	K	K^2	$\frac{K-1}{K+1}$	$\frac{K+1}{K-1}$	$\frac{K}{K^2-1}$	$\frac{K^2-1}{K}$ $= K - r$	$\frac{K^2+1}{K^2-1}$	$\frac{K-1}{K}$ $= 1 - r$	$\frac{K}{K-1}$ $= \frac{1}{1-r}$	$\frac{1}{K-1}$	**n (dB)**
0.2	0.97724	1.023292	1.047128	0.011512	86.866	21.713	0.046052	43.437	0.022762	33.933	32.933	0.2
0.5	0.94406	1.059254	1.12202	0.028774	34.754	8.6810	0.1 1519	17.391	0.055939	17.877	16.877	0.5
0.8	0.91201	1.096477	1.20227	0.046019	21.730	5.4209	0.18447	10.888	0.087988	11.365	10.365	0.8
1.0	0.89125	1.12202	1.25893	0.057502	17.391	4.3335	0.23077	8.7237	0.10875	9.1954	8.1954	1.0
1.2	0.87096	1.14815	1.31826	0.068968	14.499	3.6076	0.27719	7.2842	0.12904	7.7499	6.7499	1.2
1.4	0.85114	1.17490	1.38038	0.080418	12.435	3.0888	0.32376	6.2579	0.14886	6.7176	5.7176	1.4
1.5	0.84139	1.18850	1.41254	0.086132	11.610	2.8809	0.34711	5.8480	0.15861	6.3050	5.3050	1.5
1.8	0.81283	1.23027	1.51356	0.103249	9.6853	2.3956	0.41744	4.8944	0.18717	5.3427	4.3427	1.8
2.0	0.79433	1.25893	1.58489	0.11463	8.7241	2.1523	0.46460	4.4195	0.20567	4.8620	3.8620	2.0
2.2	0.77625	1.28825	1.65959	0.12597	7.9384	1.9531	0.51200	4.0322	0.22375	4.4692	3.4692	2.2
2.4	0.75858	1.31826	1.73780	0.13728	7.2842	1.7867	0.55968	3.7108	0.24142	4.1421	3.1421	2.4
2.5	0.74989	1.33352	1.77828	0.14293	6.9966	1.7133	0.58363	3.5698	0.25011	3.9983	2.9983	2.5
3.0	0.70795	1.41254	1.99526	0.17100	5.8480	1.4192	0.70459	3.0095	0.29205	3.4240	2.4240	3.0
3.5	0.66834	1.49623	2.2387	0.19879	5.0304	1.2079	0.82789	2.6147	0.33166	3.0152	2.0152	3.5
4.0	0.63096	1.58489	2.5119	0.22627	4.4194	1.0483	0.95393	2.3229	0.36904	2.7097	1.7097	4.0
4.5	0.59566	1.67880	2.8184	0.25340	3.9464	0.92323	1.08314	2.0999	0.40434	2.4732	1.4732	4.5
5.0	0.56234	1.77828	3.1623	0.28013	3.5698	0.82241	1.21594	1.9249	0.43766	2.2849	1.2849	5.0
5.5	0.53088	1.88365	3.5481	0.30643	3.2633	0.73922	1.35277	1.7849	0.46912	2.1317	1.1317	5.5
6.0	0.50119	1.99526	3.9811	0.33228	3.0095	0.66932	1.49407	1.6709	0.49881	2.0048	1.0048	6.0
6.5	0.47315	2.1135	4.4668	0.35764	2.7961	0.60964	1.6403	1.5769	0.52685	1.89807	0.89807	6.5
7.0	0.44668	2.2387	5.0119	0.38246	2.6146	0.55801	1.7920	1.4985	0.55332	1.80730	0.80730	7.0
7.5	0.42170	2.3714	5.6234	0.40677	2.4854	0.51291	1.9497	1.4326	0.57830	1.72918	0.72918	7.5
8.0	0.39811	2.5119	6.3096	0.43051	2.3228	0.47309	2.1138	1.3767	0.60180	1.66142	0.66142	8.0
8.5	0.37584	2.6607	7.0795	0.45366	2.2043	0.43765	2.2849	1.3290	0.62416	1.60216	0.60216	8.5
9.0	0.35481	2.8184	7.9433	0.47622	2.0999	0.40592	2.4636	1.2880	0.64519	1.54993	0.54993	9.0
9.5	0.33497	2.9854	8.9125	0.49817	2.0074	0.37730	2.6504	1.2528	0.66503	1.50368	0.50368	9.5
10.0	0.31623	3.1623	10.000	0.51950	1.9249	0.35137	2.8561	1.2222	0.68377	1.46247	0.46247	10.0
12.0	0.25119	3.9811	15.849	0.59848	1.6709	0.26811	3.7299	1.1347	0.74881	1.33545	0.33545	12.0
14.0	0.19953	5.0119	25.119	0.66733	1.4985	0.20780	4.8124	1.0829	0.80047	1.24926	0.24926	14.0
16.0	0.15849	6.3096	39.811	0.72639	1.3767	0.16257	6.1511	1.0515	0.84151	1.18834	0.18834	16.0
18.0	0.12589	7.9433	63.096	0.77637	1.2880	0.12792	7.8174	1.03220	0.87411	1.14402	0.14402	18.0
20.0	0.100000	10.0000	100.000	0.81818	1.2222	0.10101	9.9000	1.02020	0.90000	1.11111	0.11111	20.0
22:0	0.079433	12.589	158.49	0.85282	1.1726	0.079935	12.510	1.01270	0.92057	1.08629	0.086291	22.0
24.0	0.063096	15.849	251.19	0.88130	1.1347	0.063348	15.786	1.00799	0.93690	1.06734	0.067345	24.0
26.0	0.050119	19.953	398.11	0.90455	1.1055	0.050246	19.903	1.00504	0.94988	1.05276	0.052762	26.0
28.0	0.039811	25.119	630.96	0.92343	1.0829	0.039874	25.079	1.00317	0.96019	1.04146	0.041461	28.0

Table 22-1. K″ Factors for Calculating Attenuator Loss Values (Continued)

	a	b	c	d	e	f	g	h	i	j	l	
n (dB)	$r = \frac{1}{K}$	K	K^2	$\frac{K-1}{K+1}$	$\frac{K+1}{K-1}$	$\frac{K}{K^2-1}$	$\frac{K^2-1}{K}$ $= K-r$	$\frac{K^2+1}{K^2-1}$	$\frac{K-1}{K}$ $= 1-r$	$\frac{K}{K-1}$ $= \frac{1}{1-r}$	$\frac{1}{K-1}$	**n (dB)**
30.0	0.031623	31.623	1000.0	0.93869	1.0653	0.031655	31.591	1.00200	0.96836	1.03266	0.032655	30.0
32.0	0.025119	39.811	1584.9	0.95099	1.051 5	0.025135	39.786	1.00126	0.97488	1.02577	0.025766	32.0
34.0	0.019953	50.119	2511.9	0.96088	1.04072	0.019961	50.099	1.00080	0.98005	1.02036	0.020359	34.0
36.0	0.015849	63.096	3981.1	0.96880	1.03221	0.015853	63.080	1.00050	0.98415	1.01610	0.016104	36.0
38.0	0.012589	79.433	6309.6	0.97513	1.02550	0.012591	79.420	1.00032	0.98741	1.01275	0.012750	38.0
40.0	0.0100000	100.000	10,000.	0.98020	1.02020	0.0100010	99.990	1.00020	0.99000	1.01010	0.010101	40.0
42.0	0.0079433	125.89	15,849.	0.98424	1.01601	0.0079436	125.88	1.00013	0.99206	1.00801	0.0080070	42.0
44.0	0.0063096	158.49	25,119.	0.98746	1.01270	0.0063096	158.49	1.00008	0.99369	1.00635	0.0063496	44.0
46.0	0.0050119	199.53	39,811.	0.99003	1.01007	0.0050119	199.53	1.00005	0.99499	1.00504	0.0050370	46.0
48.0	0.0039811	251.19	63,096.	0.99207	1.00799	0.0039811	251.19	1.000032	0.99602	1.00400	0.0039970	48.0
50.0	0.0031623	316.23	100,000.	0.99370	1.00634	0.0031623	316.23	1.000020	0.99684	1.00317	0.0031723	50.0
60.0	0.0010000	1000.0	10^6	0.99800	1.00200	0.0010000	1000	1.000002	0.99900	1.00100	0.0010010	60.0
70.0	0.00031623	3162.3	10^7	0.99937	1.00063	0.00031623	3162.3	1.000000	0.99968	1.00032	0.00031633	70.0
80.0	0.00010000	10,000.0	10^8	0.99980	1.00020	0.00010000	10,000	1.000000	0.99990	1.00010	0.00010001	80.0
90.0	0.00003163	31,623.0	10^9	0.99994	1.00006	0.00003162	31,623	1.000000	0.99997	1.00003	0.000031624	90.0
100	0.00001000	10^5	10^{10}	0.99998	1.00002	0.00001000	10^5	1.000000	0.99999	1.00001	0.000010000	100

22.1.2 Loss

The term *loss* is constantly used in attenuator and pad design. Loss is a decrease in the power, voltage, or current at the output of a device compared to the power, voltage, or current at the input of the device. The loss in decibels may be calculated by means of one of the following equations:

$$dB_{loss} = 10 log \frac{P_1}{P_2} \tag{22-3}$$

$$dB_{loss} = 20 log \frac{V_1}{V_2} \tag{22-4}$$

$$dB_{loss} = 20 log \frac{I_1}{I_2} \tag{22-5}$$

where,

P_1 is the power at the input,

P_2 is the power at the output,

V_1 is the voltage at the input,

V_2 is the voltage at the output,

I_1 is the current at the input,

I_2 is the current at the output.

The *insertion loss* is created by the insertion of a device in an electrical circuit. The resulting loss is generally expressed in decibels.

A *minimum-loss pad* is a pad designed to match circuits of unequal impedance with a minimum loss in the matching network. This minimum loss is dependent on the ratio of the terminating impedances.

The minimum loss for attenuators of unequal impedance may be read from the graph in Fig. 22-5.

The graph is entered at the bottom at the desired impedance ratio and then followed vertically until it intersects the diagonal line. The minimum loss in decibels is then read at the left margin. For instance, assume an impedance of 600 Ω is to be matched to an impedance of 150 Ω; this is an impedance ratio of four. For this ratio, the graph indicates a minimum loss of 11.5 dB, which is the lowest value for which a passive attenuator can be designed. In actual practice the network would be designed for a loss of 12–15 dB.

22.1.3 Impedance Matching

An *impedance-matching network* is a noninductive, resistive network designed for insertion between two or more circuits of equal or unequal impedance. When properly designed, the network reflects the correct

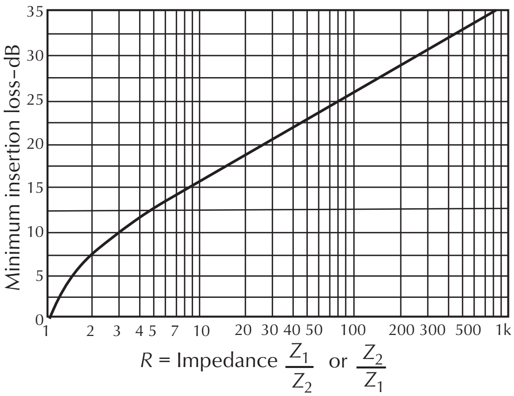


Figure 22-5. Minimum loss graph for networks of unequal impedances.

impedance to each branch of the circuit. A noninductive resistor is a resistor having little or no self-inductance.

If two resistive networks are mismatched, generally the frequency characteristics are not affected; only a loss in level occurs. If the impedance mismatch ratio is known, the loss in level may be directly read from the graph in Fig. 22-5 or with the equation

$$dB_{loss} = 20log\left(\sqrt{\frac{Z_1}{Z_2}} + \sqrt{\frac{Z_1}{Z_2} - 1}\right) \quad (22\text{-}6)$$

where,

Z_1 is the higher impedance in ohms,

Z_2 is the lower impedance in ohms.

The equation used for designing a minimum-loss attenuator when only the larger impedance Z_1 is to be matched is

$$R_1 = Z_1 - Z_2 \quad (22\text{-}7)$$

Only a series resistor *R* is used, Fig. 22-6.

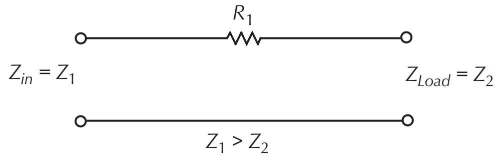


Figure 22-6. Impedance matching a low-impedance load to a high-impedance source.

If the smaller impedance is to be matched, use the following equation

$$R = \frac{Z_1 Z_2}{Z_1 - Z_2} \quad (22\text{-}8)$$

The resistor is shunted across the line, Fig. 22-7.

22.1.4 Installations, Practices, and Measurements

It is not good practice to build pads of over 40 dB loss unless special precautions are taken to reduce the distributed capacity and leakage between the input and output sections. It is more practical to build two or more pads of lower loss and connect them in tandem. The total loss is the sum of the individual losses, assuming that all impedance matches are satisfied between sections.

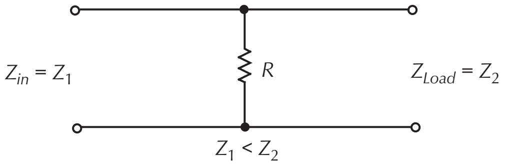


Figure 22-7. Impedance matching a high-impedance load to a low-impedance source.

When installing attenuators, the input and output circuits must be separated from each other and well shielded and grounded to prevent leakage at the higher frequencies. As an example: an attenuator of 40 dB loss has a signal voltage reduction of 100:1 between the input and output terminals. Therefore, if coupling between the input and output circuits is permitted, serious leakage can occur at frequencies above 1000 Hz.

The resistance of an attenuator can be measured with an ohmmeter by terminating the output with a resistance equal to the terminating impedance and measuring the input resistance. The resistance as measured by the ohmmeter should equal the impedance of the pad. If the attenuator is variable, the dc resistance should be the same for all steps.

If the impedance of an attenuator is not known, its value can be determined by first measuring the resistance looking into one end with the far end open and then shorted. The impedance (*Z*) is the geometric mean of the two readings

$$Z = \sqrt{Z_1 Z_2} \quad (22\text{-}9)$$

where,

Z_1 is the resistance in ohms measured with the far end open

Z_2 is the resistance in ohms measured with the far end shorted.

This measurement will hold true only for pads designed to be operated between equal terminations. If the dc resistance of the two ends differs, the pad was designed to be operated between unequal impedances.

If an attenuator is to be converted to a different impedance, the new resistors can be calculated by

$$R_x = \frac{Z_x R}{Z} \tag{22-10}$$

where,

Z_x is the new impedance in ohms,
Z is the known impedance in ohms,
R is the known value of resistance in ohms,
R_x is the new value of resistance in ohms.

Any balanced or unbalanced attenuator may be directly connected to another, provided the impedance match is satisfied and the configurations are of such nature they will not cause an unbalanced condition. Fig. 22-8A shows how an L, a bridged-T, and a plain-T pad may be connected in tandem. In Fig. 22-8B the method of connecting balanced attenuator configurations in tandem is shown.

22.2 Types of Attenuators

22.2.1 L Pads

L pads are the simplest form of attenuator and consist of two resistive elements connected in the form of an L, Fig. 22-9. This pad does not reflect the same impedance in both directions. An impedance match is afforded only in the direction of the arrow shown in the figures. If an L-type network is employed in a circuit that is sensitive to impedance match, the circuit characteristics may be affected. An L-type network should not be used, except where a minimum loss is required and a network of the T configuration will not serve because its minimum loss is too high.

For unequal impedances, the impedance match may be in the direction of the larger or the smaller impedance but not both.

If the network is designed to match the impedance in the direction of the series arm, the mismatch is toward the shunt arm. The mismatch increases with the increase of loss, and, at high values of attenuation, the value of the shunt resistor may become a fraction of an ohm, which can have a serious effect on the circuit to which it is connected.

The configuration for an L-type network operating between impedances of unequal value, Z_1 and Z_2, is shown in Fig. 22-9A. The impedance match is toward the larger of the two impedances, Z_1, and the values of the resistors are

$$R_1 = \frac{Z_1}{S}\left(\frac{KS-1}{K}\right) \tag{22-11}$$

$$R_2 = \frac{Z_1}{S}\left(\frac{1}{K-S}\right) \tag{22-12}$$

where,

S is $\sqrt{\frac{Z_1}{Z_2}}$,

The value of K is taken from Table 22-1.

For a condition where the impedances are equal and the impedance match is in the direction of the arrows, Fig. 22-9B, the values of the resistors may be calculated by the equation:

$$R_1 = Z(i) \tag{22-13}$$

$$R_2 = Z(l) \tag{22-14}$$

The values of *i* and *l are* taken from Table 22-1.

When the impedances are unequal and the impedance match is toward the smaller of the two impedances, Fig. 22-9C, the values of the resistors are determined by the equations

$$R_1 = \frac{Z_1}{S}(K-S) \tag{22-15}$$

$$R_2 = \frac{Z_1}{S}\left(\frac{K}{KS-1}\right) \tag{22-16}$$

where,

S is $\sqrt{\frac{Z_1}{Z_2}}$.

For the conditions shown in Fig. 22-9D, resistors R_1 and R_2 are calculated by

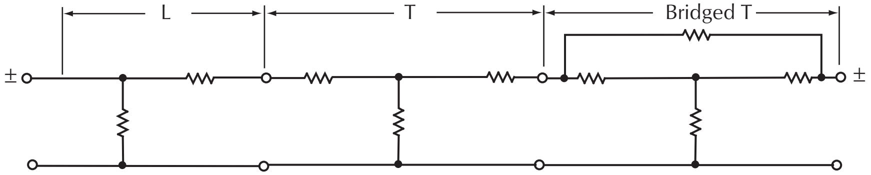
$$R_1 = Z(K-1) \tag{22-17}$$

$$R_2 = Z(l) \tag{22-18}$$

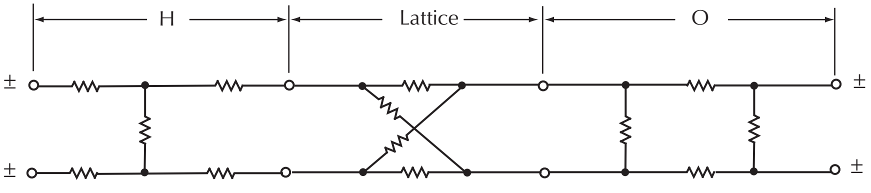
The values of K and l are taken from Table 22-1.

If a minimum-loss, L attenuator is used to match two impedances of unequal value, as in Fig. 22-9A, the resistor values will be

$$R_1 = \sqrt{Z_1(Z_1-Z_2)} \tag{22-19}$$

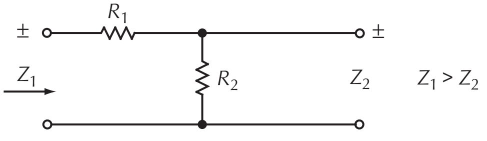


A. Unbalanced

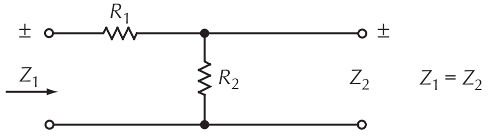


B. Balanced

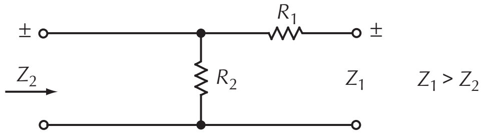
Figure 22-8. Attenuators connected in tandem.



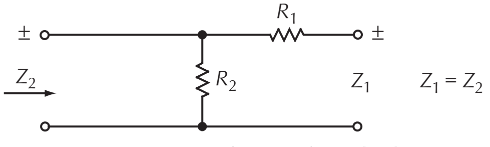
A. Between impedances of unequal value.



B. Between impedances of equal value.



C. Impedance unequal and impedance match toward the smaller of the two.



D. Between impedances of equal value in the direction of the shunt arm.

Figure 22-9. Configurations of L-type networks.

$$R_2 = \frac{Z_1 Z_2}{R_1} \tag{22-20}$$

where,

R_1 is the series resistor in ohms connected on the side of the larger impedance;

R_2 is the shunt resistor in ohms.

The loss through the attenuator will be

$$dB_{loss} = 20 log\left(\sqrt{\frac{Z_1}{Z_2}} + \sqrt{\frac{Z_1}{Z_2} - 1}\right) \tag{22-21}$$

22.2.2 Dividing Networks

Dividing or *combining networks* are resistive networks designed to combine several devices or circuits, each having the same impedance, Fig. 22-10A. The resistors may be calculated with the equation

$$R_B = \frac{N-1}{N+1}Z \tag{22-22}$$

where,

R_B is the build-out resistor in ohms,

N is the number of circuits fed by the source impedance,

Z is the circuit impedance in ohms.

The loss of the network is

$$dB_{loss} = 20 log(N-1) \tag{22-23}$$

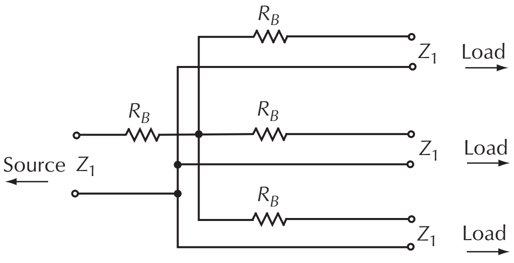
where,

N is the number of input or output circuits.

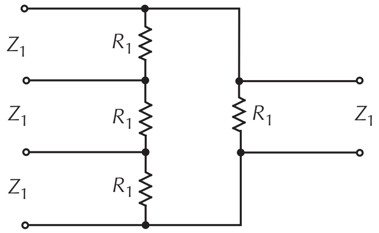
Unused circuits of a dividing or combining network must be terminated in a resistive load equal to the normal load impedance.

This same circuit can be reversed and used as a combining network. This circuit was often used in the design of sound mixers.

Combining or branching networks may also be designed as a series configuration, Fig. 22-10B. For equal impedances the equation is



A. Combining or dividing network for matching a single circuit to three others.



B. Series combining network for combining one circuit to three others.

Figure 22-10. A combining or dividing network for matching a single circuit to three circuits.

$$R_1 = \frac{N-1}{N+1}Z \qquad (22\text{-}24)$$

where,
R_1 is the terminating resistor in ohms,
N is the number of branch circuits,
Z is the circuit impedance in ohms.

The insertion loss may be calculated:

$$dB_{loss} = 20log(N-1) \qquad (22\text{-}25)$$

where,
N is the number of branch circuits.

A series configuration can only be used in an ungrounded circuit. The insertion loss of a combining network may be avoided by the use of an active combining network (see Sections 22.2.15 and 22.2.16).

22.2.3 T Attenuators

A *T-type attenuator* is an attenuator network consisting of three resistors connected in the form of a T, Fig. 22-1. The network may be designed to supply an impedance match between circuits of equal or unequal impedance. When designed for use between circuits of unequal impedance, it is often referred to as a *taper pad*.

If a T pad is to work between equal impedances, the resistor values will be

$$\begin{aligned} R_1 &= R_2 \\ &= Z(d) \end{aligned} \qquad (22\text{-}26)$$

$$R_3 = 2Z(f) \qquad (22\text{-}27)$$

where,
Z is the input and output impedance in ohms,
R_1 and R_2 are the series resistors in ohms,
R_3 is the shunt arm in ohms,
The values of d and f are taken from Table 22-1.

A T type attenuator may be designed for any value of loss if designed to operate between equal impedances.

The resistors for a T pad of unequal impedances are calculated with the following equations:

$$R_1 = Z_1h - 2\sqrt{Z_1Z_2}(f) \qquad (22\text{-}28)$$

$$R_2 = Z_2h - 2\sqrt{Z_1Z_2}(f) \qquad (22\text{-}29)$$

$$R_3 = 2\sqrt{Z_1Z_2}(f) \qquad (22\text{-}30)$$

where,
Z_1 is the larger of the two impedances.
The values of f and h are taken from Table 22-1.

Thus, for a network to match 600 Ω to a circuit of 250 Ω with a loss of 20 dB, the resistor values are

$$\begin{aligned} R_1 &= 600(1.0202) - 2\sqrt{150,000}(0.10101) \\ &= 533.88\ \Omega \\ R_2 &= 250(1.0202) - 2\sqrt{150,000}(0.10101) \\ &= 176.81\ \Omega \\ R_3 &= 2\sqrt{150,000}(0.10101) \\ &= 78.7\ \Omega. \end{aligned}$$

A *balanced T pad* is called an *H pad*. The pad is first calculated as an unbalanced T configuration. The series resistance elements are then divided and one-half connected in each side of the line, Fig. 22-2. The shunt resistor remains the same value as for the unbalanced configuration. A tap is placed at the exact electrical center of the shunt resistor for connection to ground.

The average noise level for a T pad is −100 dB and constant. Therefore, the signal-to-noise level varies with the amount of attenuation.

22.2.4 Bridged T Attenuators

A *bridged T pad* is an attenuator network containing four resistive elements, Fig. 22-11. The resistors are equal in value to the line impedance; therefore, they require no calculation. This network is designed to work between impedances of equal value only. The contact arms for resistors R_5 and R_6 are connected mechanically by a common shaft and vary inversely in value with respect to each other.

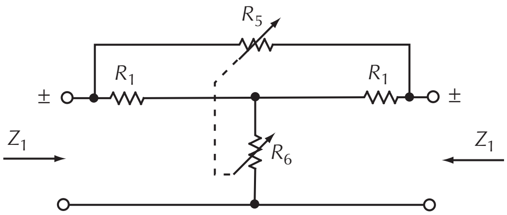


Figure 22-11. A bridged T attenuator. For variable pads, the arms R_5 and R_6 are made variable.

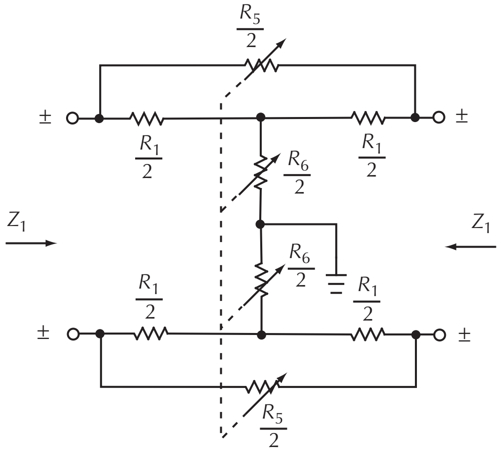


Figure 22-12. Balanced bridged T attenuator. For a variable configuration, variable arms are required.

A *balanced bridged T attenuator* is a configuration similar to the unbalanced bridged T attenuator, except the resistor elements are divided and placed in each side of the line, Fig. 22-12. The principal objection to the use of this configuration, if made variable, is that the shunt resistor R_6 must be divided into two separate arms to provide a ground connection at the exact electrical center. However, if the circuit feeding or terminating the attenuator is balanced to the ground, the ground connection at the attenuator center will not be required.

The resistor values are calculated with the following equations:

$$R_1 = Z \tag{22-31}$$

$$R_5 = (K-1)Z \tag{22-32}$$

$$R_6 = Z(l) \tag{22-33}$$

where,
Z is the line impedance in ohms,
R_5 is the bridging resistor in ohms,
R_6 is the shunt resistor in ohms.
The values of K and l are taken from Table 22-1.

The impedance variations for a typical high quality attenuator used in a mixer network are shown in Fig. 22-13. The greatest impedance variation occurs as the attenuator arm approaches zero attenuation and amounts to about 80 Ω. This impedance variation is not too serious, as the mixer-combining network with its building-out resistors isolates this variation to a great extent from associated attenuators.

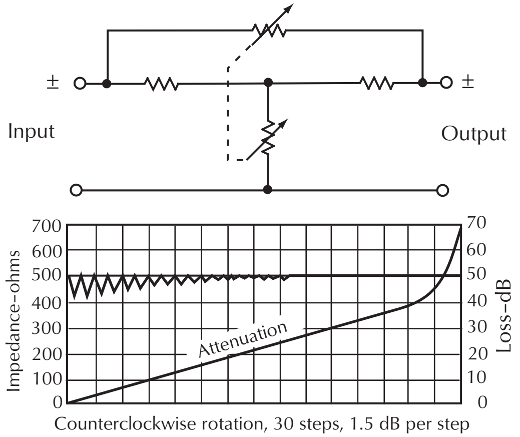


Figure 22-13. Impedance characteristics for a high quality variable bridged T attenuator.

22.2.5 π or Δ Attenuators

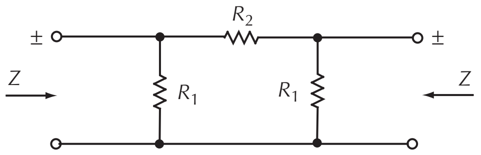
A π or Δ *attenuator* is a resistive network resembling the Greek letter pi (π), or delta (Δ), Fig. 22-14. Such networks may be used between impedances of equal or unequal values.

For networks operating between impedances of equal value

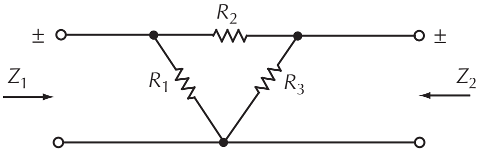
$$R_1 = Z(e) \tag{22-34}$$

$$R_2 = \frac{Z}{2}(g) \tag{22-35}$$

where,
R_1 is the input and output resistor in ohms,
R_2 is the series resistor in ohms,
Z is the input and output impedance in ohms,
Find e and g in Table 22-1.



A. Between impedances of equal value.



B. Between impedances of unequal value.

Figure 22-14. π and Δ attenuators.

When the impedances are unequal values, the resistors are calculated with the following equations:

$$R_1 = Z_1\left(\frac{K^2-1}{K^2-2KS+1}\right) \tag{22-36}$$

$$R_2 = \frac{\sqrt{Z_1 Z_2}}{2}(g) \tag{22-37}$$

$$R_3 = Z_2\left(\frac{K^2-1}{K^2-2\frac{K}{S}+1}\right) \tag{22-38}$$

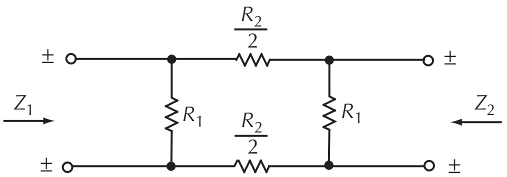
where,
R_1 and R_3 are shunt resistors in ohms,
R_2 is the series resistor in ohms,
Z_1 is the input impedance in ohms,
Z_2 is the output impedance in ohms,
S is $\sqrt{\frac{Z_1}{Z_2}}$.

The values of K, K^2, and g are taken from Table 22-1.

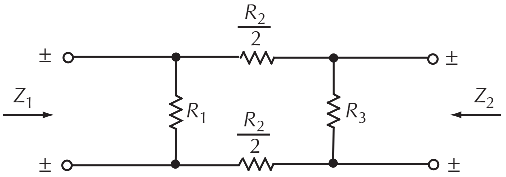
An O-type attenuator is a balanced π attenuator. The circuit element values may be obtained by first calculating for a π configuration and then dividing the series resistor and placing half in each side of the line, Fig. 22-15. The shunt resistors remain the same value.

22.2.6 U Attenuators

U attenuators, Fig. 22-16, may be of a symmetrical or balanced-type configuration and are useful for matching a high impedance to a low impedance. The impedance match is of first importance, the loss being secondary.



A. Between impedances of equal value.



B. Between impedances of unequal value.

Figure 22-15. Balanced π or O attenuator.

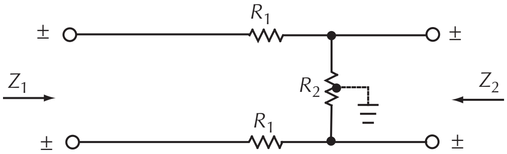


Figure 22-16. A U pad configuration for operation between impedances of unequal value.

For a symmetrical configuration to work between unequal impedances when the impedance match of Z_1 is important, the resistors may be calculated as follows:

$$R_1 = \frac{Z_1}{2S}\left(\frac{KS-1}{K}\right) \tag{22-39}$$

$$R_2 = \frac{Z_1}{S}\left(\frac{1}{K-S}\right) \tag{22-40}$$

where,
R_1 is the series resistor in ohms,
R_2 is the shunt resistor in ohms,
Z_1 is the larger impedance in ohms,
Z_2 is the smaller impedance in ohms,
S is $\sqrt{\frac{Z_1}{Z_2}}$,

The value of K is taken from Table 22-1.

When the low impedance, Z_2 is to be matched the equations are:

$$R_1 = \frac{Z_1}{2S}(K-S) \tag{22-41}$$

$$R_2 = \frac{Z_1}{S}\left(\frac{K}{KS-1}\right) \tag{22-42}$$

where,

S is $\sqrt{\dfrac{Z_1}{Z_2}}$,

The value of K is taken from Table 22-1.

Any U pad may be balanced to ground by connecting a ground to the electrical center of the shunt resistor.

22.2.7 Ladder Attenuators

Ladder-type pads, Fig. 22-17, are so named because they look like a ladder laying on its side. The ladder pad is actually a group of *pi* attenuators in tandem, R_2 being common to each section. Because of resistor R_4, this type of attenuator has a fixed 6 dB loss, exclusive of the attenuator setting, which must be taken into account when designing a ladder attenuator. The ladder attenuator does not have a constant input and output impedance throughout its range of attenuation. However it does reflect a stable impedance into its source.

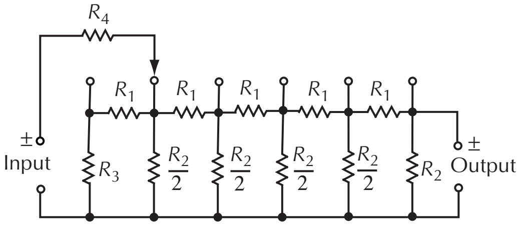


Figure 22-17. Unbalanced ladder attenuator with five fixed steps of loss.

Ladder potentiometers for mixercontrol use may be obtained in two types of construction—slide-wire and contact types.

For mixers, the slide-wire type control is generally employed because it permits a smooth, even attenuation over a wide range. The contact type, although not quite as smooth in operation as the slide-wire, has only one row of contacts, which reduces the noise and maintenance.

Ladder networks may also be designed for balanced operation. This is accomplished by connecting two unbalanced networks side by side, Fig. 22-18. The circuit elements are not divided in the same manner as for other types balanced networks. If an unbalanced ladder network is compared with a balanced ladder network, resistors R_1 are divided by two, resistors R_4 are also divided by two, and at the output R_2 is now twice the value for the unbalanced configuration. Resistor R_3 remains at its original value on each side of ground.

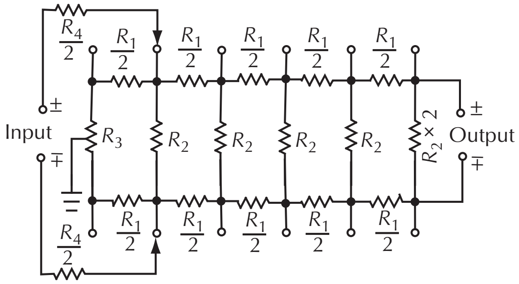


Figure 22-18. Balanced ladder attenuator.

The equations used to calculate the resistor values are:

$$R_1 = \frac{K^2-1}{2K}(Z) \tag{22-43}$$

$$R_2 = Z(e) \tag{22-44}$$

$$R_3 = \frac{R_2 Z}{R_2 + Z} \tag{22-45}$$

$$R_4 = \frac{Z}{2} \tag{22-46}$$

where,
R_1 is the series resistance in ohms,
R_2 is the shunt resistance in ohms,
R_3 is the input shunt resistor in ohms,
R_4 is the series resistance in the contact arm circuit in ohms,
The values of K and K^2 are taken from Table 22-1.

The value of K is dependent on the loss per step, not the total loss.

The noise level for a ladder attenuator is on the order of –120 dB, and as the attenuation increases, the SNR increases. This type of attenuator will show impedance variations at both the input and output and between steps. However, when used in a combining network with the proper building-out resistors, these variations are of little consequence. A typical impedance curve is shown in Fig. 22-19.

22.2.8 Simple Volume and Loudness Controls

A simple volume control consists of a potentiometer with the two ends connected to the source and the wiper

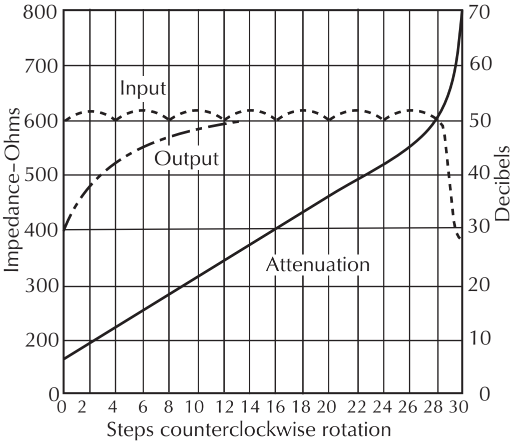


Figure 22-19. Impedance characteristics of a 600 Ω ladder type attenuator.

and one end connected to the load, Fig. 22-20. The volume control should be a high impedance with respect to the source so it will not load it, and the load impedance should be a high enough so as not to affect the control. The output voltage is calculated with the following equation:

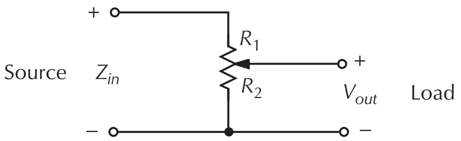

+
−
Source
Z_{in}
R_1
R_2
+
−
V_{out}
Load

Figure 22-20. Simple volume control.

$$V_{out} = V_{in} \frac{\left(\frac{R_2 Z_2}{R_2 + Z_2}\right)}{R_1 + \left(\frac{R_2 Z_2}{R_2 + Z_2}\right)} \tag{22-47}$$

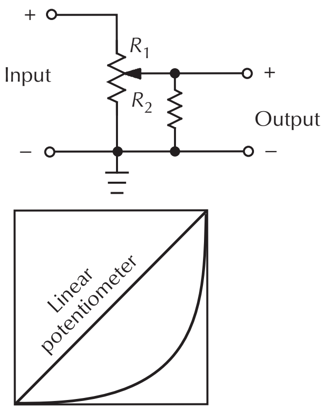
where,

R_1 is the upper section of control,
R_2 is the lower section of control,
Z_2 is the load impedance.

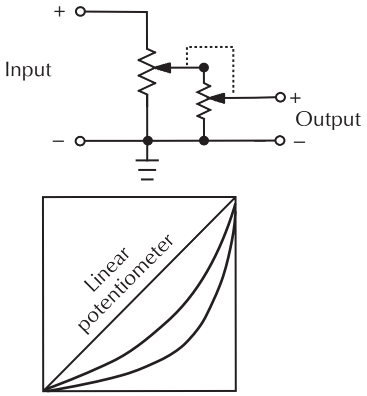
If the load impedance is high compared to R_2, the equation is simplified to

$$V_{out} = V_{in}\left(\frac{R_2}{R_1 + R_2}\right) \tag{22-48}$$

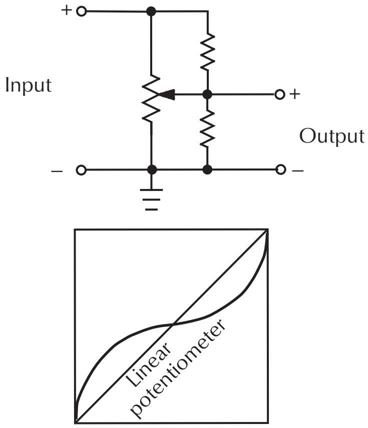
The attenuation is


+
−
Input
R_1
R_2
+
−
Output
Linear potentiometer

A. Shunt connected from the wiper to the ground.


+
−
Input
+
−
Output
Linear potentiometer

B. Second potentiometer ganged with the straight-line potentiometer.


+
−
Input
+
−
Output
Linear potentiometer

C. Two shunt resistors connected at each side of the wiper.

Figure 22-21. Method of varying the response of a simple potentiometer.

$$dB = 10 log(4)\left[\frac{R_1 + \left(\frac{R_2 Z_2}{R_2 + Z_2}\right)^2}{Z_1 Z_2}\right] \tag{22-49}$$

Normally, volume controls have a logarithmic taper, so the first 50% of the pot only represents a change of 7–8%, following the ear's sensitivity. If a special taper is required, a linear pot can be altered to change its characteristics by shunting a fixed resistor from one end of the potentiometer to the wiper. Three methods of shunting a straight-line potentiometer are shown in Fig. 22-21. In the first method, the shunt resistor is connected from the wiper to ground. With the correct value shunt resistance, the potentiometer will have a taper relative to the angular rotation, as shown below the schematic. The second method makes use of a second potentiometer ganged with the straight-line potentiometer. In the third method, two shunt resistors connected at each side of the wiper result in a taper resembling a sine wave. A fourth method, not shown, uses a shunt resistor connected from the wiper to the top of the potentiometer.

A loudness control incorporates a circuit to alter the frequency response to follow the Fletcher-Munson curves of equal loudness—i.e., the softer the level, the more the low frequencies must be boosted with respect to 1 kHz and above. To approximate this a capacitor is tapped off the volume control at about 50% rotation. As the wiper is rotated below the tap, the signal has the high frequencies rolled off, giving the effect of low-frequency boost.

22.2.9 Light-Dependent Attenuators

A *light-dependent attenuator* (LDA) is one where the attenuation is controlled by varying the intensity of a light source on a light-dependent resistor *(*LDR) (cadmium sulfide cell). LDAs were popular before op-amps and are still useful for remote control as they are not affected by noise or hum on the control line. LDAs eliminate problems of noisy potentiometers as the potentiometers operate the lamp circuit that has an inherent lag time. This type of circuit is also very useful for remote control as the remote control line carries lamp control voltage so it is not susceptible to hum and extraneous pickup.

A simple volume control is shown in Fig. 22-22. *R* and LDR form an attenuator. When the light source is bright, the resistance of the LDR is low; therefore, most of the signal is dropped across *R*. When the light intensity is decreased, the resistance of LDR increases and more signal appears across the LDR. This circuit has constantly varying impedances.

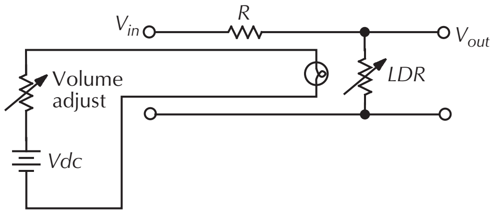


Figure 14-22. Volume control using a light-dependent resistor.

A constant impedance attenuator would require more LDRs and light sources to approximate a constant impedance type of attenuator.

The advantages of a LDA are:

1. No wiper noise.
2. One control can operate many attenuators.
3. Controls can be remoted from the attenuator.

The disadvantages are:

1. Lamp burnout or aging.
2. Slow response time.

22.2.10 Feedback-Type Volume Control[1]

In a *feedback-type volume control* attenuation is controlled by the amount of feedback in the circuit. Feedback-type volume controls have the advantage of reduced hum and noise as they reduce the gain of the active network rather than reducing just the signal level.

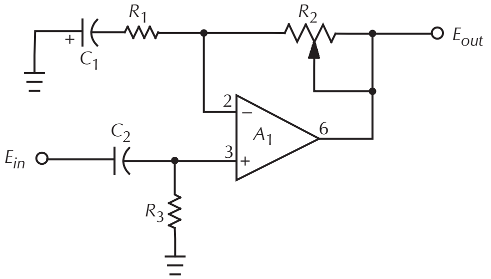


Figure 22-23. Noninverting linear-feedback, gain-controlled amplifier.

A noninverting op-amp feedback gain controlled amplifier is shown in Fig. 22-23. Feedback resistor R_2 is used to adjust the gain of the op-amp and therefore the output. When R_2 is zero, gain will be one as the system has 100% feedback. Increasing the value of R_2 decreases feedback, consequently increasing gain by the ratio of R_2/R_1. Gain can be determined with the equation

$$E_0 = E_{in}\left(\frac{R_1 + R_2}{R_1}\right) \tag{22-50}$$

22.2.11 Voltage-Controlled Amplifiers

A *voltage-controlled amplifier (VCA)* is used as an attenuator by varying a dc control voltage. VCAs are often used for automatic mixing since the control voltage can be stored in analog or digital form and on command can be programmed back into the console and VCA.

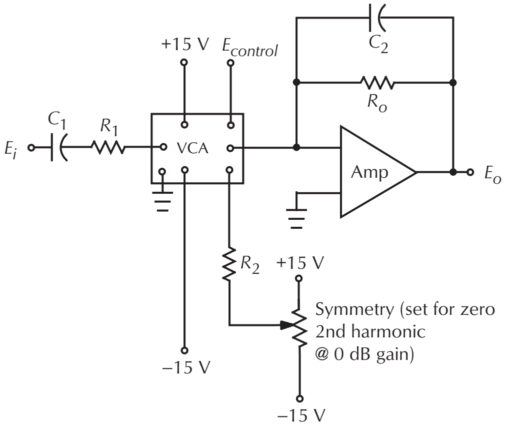


Figure 22-24. VCA volume control.

VCAs are also useful for remote control operation and in compressors or expanders. VCAs have attenuation ranges from 0–130 dB and response time better than 100 µs. A typical circuit is shown in Fig. 22-24. Since the input is a virtual-ground summing point, R_1 is used so as not to load the preceding circuit. The output circuit must feed a virtual ground so an operational amplifier current-to-voltage converter (any operational amplifier with a resistor from output to inverting input and with the noninverting input grounded) must be used. The circuit can be used with a linear taper potentiometer to give a linear control characteristic.

22.2.12 Field Effect Transistor Attenuators

A field effect transistor attenuator is one where an FET is used to control gain. Field effect transistors have characteristics much like a tube—that is, high-input impedance and moderate-output impedance. In its simplest form, the FET is used as the lower leg of a voltage divider, Fig. 22-25A.

The voltage out is

$$\begin{aligned} V_{out} &= V_{in} r_{DS(on)} + V_{out(max)} \\ &= \frac{V_{in}}{R + r_{DS(on)}} \end{aligned} \tag{22-51}$$

where,

r_{DS} is the resistance of the drain to source.

To improve distortion and linearity, feedback is required around the FET as in Fig. 22-25B. If a low-output impedance is required, an op-amp can be used in conjunction with the FET, Fig. 22-25C. In this circuit, the op-amp is used to match impedances. The FET can also be used to control feedback, Fig. 22-25D. The gain in this circuit is

$$AV = 1 + \frac{R_f}{r_{DS}} \tag{22-52}$$

where,

R_f is the feedback resistor.

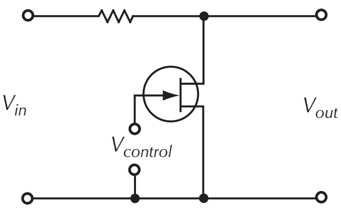
When r_{DS} is minimum, gain is maximum as most of the feedback is shorted to ground. The FET can also be used as a T attenuator, Fig. 22-25C. This provides optimum dynamic linear range attenuation and tends to hold the impedances more even.

22.2.13 Automated Faders

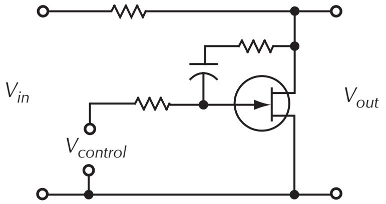
In an *automated fader*, the fade control can be programmed into a data storage device and used to adjust the fader settings during mixdown, Fig. 22-26. The fader is adjusted manually, and when the desired setting is made, a write voltage is injected into the programmer (encoder) that supplies data to the data track of the tape recorder. During playback, the data track is decoded and, through the read control, adjusts the attenuator to the recorded level. If the mixdown is not proper, any control can be adjusted or updated and the tape played over again.

22.2.14 Automatic Attenuators

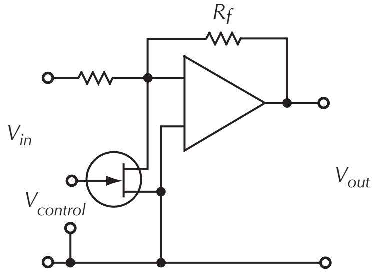
In a *automatic attenuator*, the attenuation varies automatically between two points, usually off and a prescribed setting. Automatic attenuators are often voice operated but can be manually operated. They are used to automatically turn off unused inputs, as a Ducker and gating.


V_{in}
$V_{control}$
V_{out}

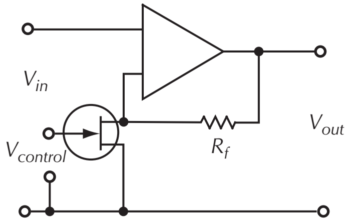
A. FET as the lower leg of a voltage divider.


V_{in}
$V_{control}$
V_{out}

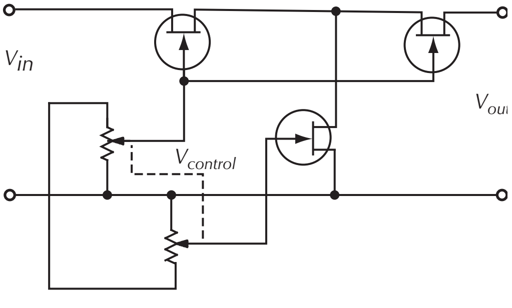
B. Feedback required around the FET.


R_f
V_{in}
$V_{control}$
V_{out}

C. An op amp used in conjunction with an FET.


V_{in}
$V_{control}$
R_f
V_{out}

D. An FET used to control feedback.


V_{in}
$V_{control}$
V_{out}

E. An FET as a T attenuator.
Figure 22-25. FET attenuators.

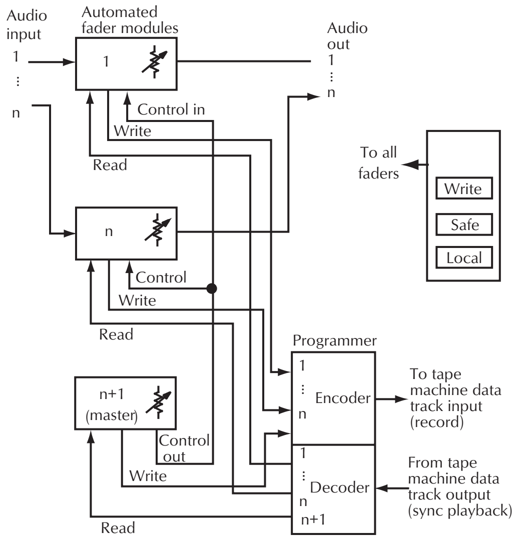

Audio input
1
⋮
n
Automated fader modules
1
Control in
Write
Read
n
Control
Write
Read
n+1 (master)
Control out
Write
Read
Audio out
1
⋮
n
To all faders
Write
Safe
Local
Programmer
1
⋮
n
Encoder
1
⋮
n
n+1
Decoder
To tape machine data track input (record)
From tape machine data track output (sync playback)

Figure 22-26. Functional block diagram of a automated fader.

22.2.15 Mixers

A mixer is a device used to mix two or more signals into one composite signal. Mixers may be adjustable or nonadjustable and either active or passive.

A passive mixer uses only passive devices (i.e., resistors and potentiometers), Fig. 22-27.

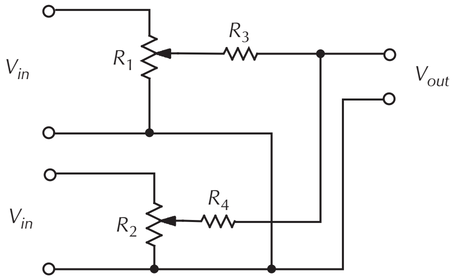

V_{in}
R_1
R_3
V_{out}
V_{in}
R_2
R_4

Figure 22-27. A passive mixer circuit.

The main disadvantage of passive mixing is that an amplifier is required after mixing to boost the gain back to the level at the input of the mixer. As the attenuator controls are lowered, the signal on the mixing buzz is reduced; however, the mixing buzz noise remains the same, so the SNR is reduced, causing more apparent noise at low levels where high signal-to-noise is most important. This can be seen in the analysis of Fig. 22-28.

In Fig. 22-28A, the input signal of –110 dBm is not attenuated; therefore, the signal going into the booster

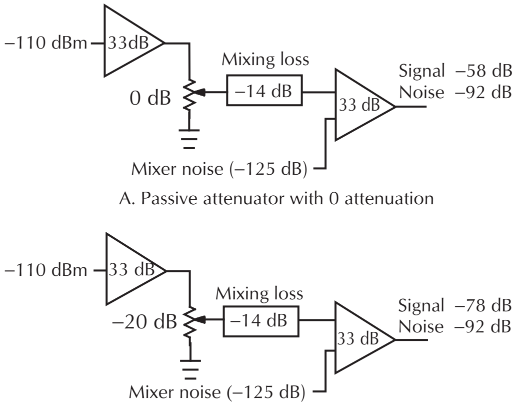


Figure 22-28. Signal-to-noise analyses of passive attenuators.

amplifier is –91 dB and out of the booster amplifier is –58 dB [–110 + (+33) + (–14) + (+33)]. The mixer noise going into the booster amplifier is –125 dB; therefore, the output noise is –92 dB [–125 + (+33)] or 34 dB below the signal.

In Fig. 22-28B, the input signal of –110 dBm is attenuated 20 dB in the mixer so the signal to the booster is –111 dB and the signal output is –78 dB. The mixer input noise is still –125 dB into the booster and –92 dB out of the booster, a difference between the signal and the noise of only 14 dB, hardly enough to be useful.

An active mixer is one that uses operational amplifiers (op-amps) or some other active device along with resistors and/or potentiometers to control gain or attenuation.

A unity-gain current-summing amplifier is used for a standard active mixer. The mixer is usually designed for an input impedance of about 5–10 kΩ, an output impedance of less than 200 Ω, and a gain of 0 to 50. A typical active mixer is shown in Fig. 22-29.

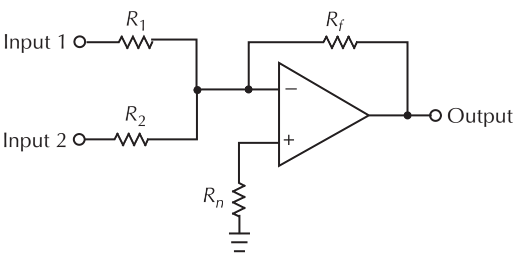


Figure 22-29. An active mixer block diagram.

In unity-gain current-summing amplifiers feedback to the minus or inverting input presents an extremely low apparent input impedance or virtual ground on the inverting input.

The positive input is also essentially ground since the current through R_n will only produce about 0.5 mV. While the positive input can be grounded, it is better to make the R_n a value about the same as the parallel combination of $R_1 + R_2 + R_f$ to reduce offset voltage.

Any small, positive-going input applied to the input of R_1 is amplified by the high-gain op-amp driving the output negative since the input signal is on the inverting input. The output signal is fed back through R_f, the feedback resistor, and it continuously attempts to drive the voltage on the input to ground.

Since the input is a virtual ground, the input impedances are determined by R_1 and R_2. The gain of the circuit is

$$\text{input } 1_{gain} = \frac{R_f}{R_1} \tag{22-53}$$

$$\text{input } 2_{gain} = \frac{R_f}{R_2} \tag{22-54}$$

If the gain of both inputs were to be the same, R_1 and R_2 would remain constant and R_f would be varied. Mixers, however, usually require separate gain control for each input so R_1 and R_2 are varied to change the gain of the system. Increasing R_1 or R_2 decreases the gain. The main disadvantage of this system is that the input impedance varies with gain.

The advantage of an active mixer is that gain is included in the mixing circuit; therefore, it does not need a gain makeup amplifier that amplifies both the signal and the mixing noise after the mixer. With active mixing, the mixing noise is also reduced along with the signal, improving the SNR, particularly at low level.

22.2.16 Summing Amplifiers[1]

A standard audio circuit function is the linear combination of a number of individual signals into a common output without crosstalk or loss. This function is well suited for the summing amplifier, which is often referred to as *active combining network*. Summing amplifiers operate much like the mixer in Section 22.2.15. Fig. 22-30 shows a 10-input summing amplifier using one op-amp. Channel isolation is important in summing amplifiers to eliminate crosstalk.

The primary determinant of interchannel isolation is the nonzero summing-bus impedance presented by the virtual ground of the inverter and, to a lesser extent, by

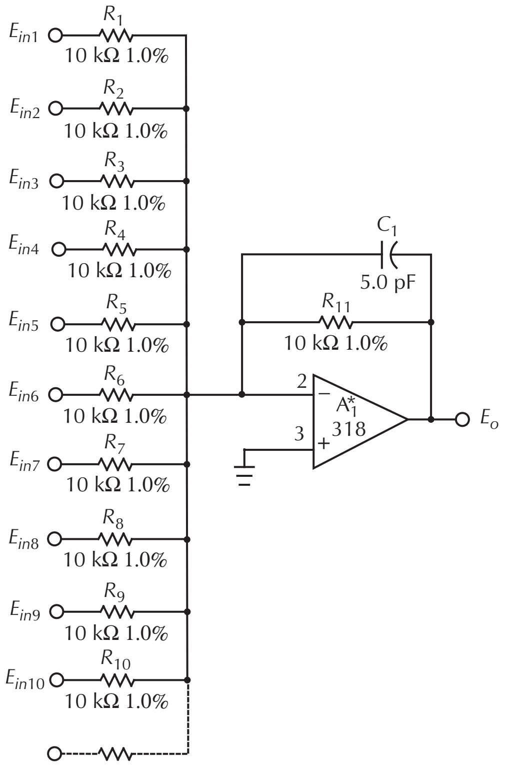

E_{in1}
R_1
10 kΩ 1.0%
E_{in2}
R_2
10 kΩ 1.0%
E_{in3}
R_3
10 kΩ 1.0%
E_{in4}
R_4
10 kΩ 1.0%
E_{in5}
R_5
10 kΩ 1.0%
E_{in6}
R_6
10 kΩ 1.0%
E_{in7}
R_7
10 kΩ 1.0%
E_{in8}
R_8
10 kΩ 1.0%
E_{in9}
R_9
10 kΩ 1.0%
E_{in10}
R_{10}
10 kΩ 1.0%
C_1
5.0 pF
R_{11}
10 kΩ 1.0%
2
3
−
+
A_1^*
318
E_o

Figure 22-30. Summing amplifier (active combining network).

the source impedances at the inputs. To illustrate the method of calculating interchannel isolation, refer to Fig. 22-31. There are two attenuations that a signal must undergo in order to leak from one channel to an adjacent channel. The first attenuation consists of R_i and R_{in}; the second consists of R_i and R_s.

The equation for calculating isolation is

$$\text{Isolation from } E_{in_a} \text{ to } E_{in_b} = \left(\frac{R_{i1}+R_{in}}{R_{in}}\right)\frac{R_{i2}+R_{sb}}{R_{in}}$$

where,

R_{sb} is the E_{sb} source resistance in ohms,

R_{in} is the A_1 closed-loop input impedance in ohms or

$$R_{in} \cong \frac{R_f}{A_{vo}\beta}$$

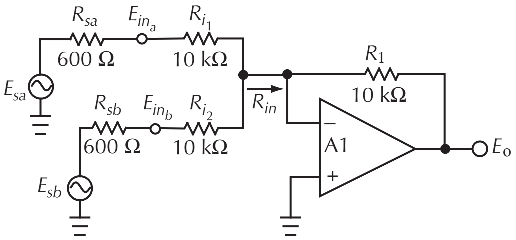

R_{sa}
E_{in_a}
R_{i_1}
600 Ω
10 kΩ
E_{sa}
R_{sb}
E_{in_b}
R_{i_2}
600 Ω
10 kΩ
E_{sb}
R_{in}
R_1
10 kΩ
A1
−
+
E_o

Figure 22-31. Method of calculating interchannel isolation for a summing amplifier.

Reference

1. Jung, Walter G. *Audio IC Op-Amp Applications,* Third Edition, Carmel, IN: SAMS, a Division of Macmillan Computer Publishing, 1989.

Chapter 23

Filters and Equalizers

by Steven McManus

23.1 Filter and Equalizer Definitions 785
  23.1.1 Pass Band 785
  23.1.2 Stop Band 785
  23.1.3 Cutoff Frequency. 785
  23.1.4 Corner Frequency 785
  23.1.5 Bandwidth 785
  23.1.6 Transition Band 785
  23.1.7 Center Frequency. 785
    23.1.7.1 Geometric Symmetry 785
    23.1.7.2 Arithmetic Symmetry 785
  23.1.8 Order 786
  23.1.9 Phase Angle 786
  23.1.10 Phase Delay 786
  23.1.11 Group Delay 786
  23.1.12 Transient Response 786
  23.1.13 Minimum Phase. 786
23.2 Passive Filters 786
  23.2.1 First-Order L and C Networks. 786
    23.2.1.1 Capacitive Networks 787
    23.2.1.2 Inductive Networks 788
  23.2.2 Second-Order L-Type Networks 788
  23.2.3 T and π Networks 789
    23.2.3.1 Low Pass 789
    23.2.3.2 High Pass 789
    23.2.3.3 Parallel Resonant Elements 790
    23.2.3.4 Series Resonant Elements 790
    23.2.3.5 Bandpass 791
    23.2.3.6 Band Reject 791
    23.2.3.7 Ladder Networks 791
  23.2.4 Filter Design 791
    23.2.4.1 Butterworth 792
    23.2.4.2 Linkwitz-Riley 792
    23.2.4.3 Chebyshev I and II 792
    23.2.4.4 Elliptical 793
    23.2.4.5 Normalizing 793
    23.2.4.6 Scaling 793
  23.2.5 Q and Damping Factor 793
  23.2.6 Impedance Matching 793
23.3 Active Filters 793

23.3.1 Filter Topologies 794
23.3.1.1 Sallen-Key 794
23.3.1.2 State Variable 795
23.3.1.3 All-Pass Filter 796
23.3.2 Pole-Zero Analysis 796
23.3.2.1 Zeros 797
23.3.2.2 Poles 797
23.3.2.3 Stability 797
23.4 Switched Capacitor Filters 797
23.5 Digital Filters 797
23.5.1 FIR Filters 798
23.5.1.1 FIR Coefficients 798
23.5.1.2 FIR Length 799
23.5.2 IIR Filters 799
23.5.2.1 Calculation of Coefficients from Poles and Zeros 799
23.6 Equalizers 800
23.6.1 Tone Control 800
23.6.2 Graphic Equalizers 800
23.6.2.1 Transversal Equalizers 801
23.6.3 Parametric Equalizers 801
23.6.3.1 Semi-Parametric Equalizers 801
23.6.3.2 Symmetric or Asymmetric Q 802
23.6.4 Programmable Equalizers 802
23.6.5 Adaptive Equalizers 802
References 802

23.1 Filter and Equalizer Definitions

A *filter* is a device or network for separating signals on the basis of their frequency. Filters can either be defined in terms of their *pass band* only where, the frequencies of interest are allowed through, or in term of their *stop band,* where certain frequencies are removed. The default design mode for most filters is as a *low pass* where all frequencies below a *cutoff* frequency, and extending down to dc, are allowed to pass. A simple re-arrangement usually allows for a *high pass* to be made, where all frequencies above a *cutoff* frequency, and extending upward, are transmitted. Other mode complex responses such as *bandpass* are constructed from these basic elements.

Passive filters have no amplification components in the circuit. They cannot add energy to the signal so can only act to attenuate signals.

Active filters use transistor or operational amplifier-based gain stages allowing the option of boosting some of the, or the whole, spectrum.

An *equalizer* is a device that uses filters to compensate for undesirable magnitude or phase characteristics of a systems response.

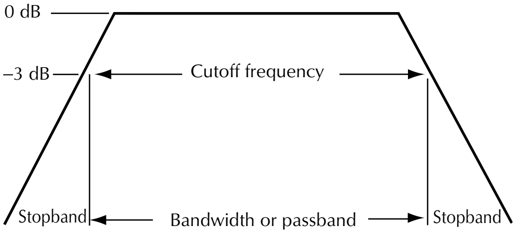


Figure 23-1. Pass bands and Stop bands of a filter.

23.1.1 Pass Band

The *pass band* is a band of frequencies that pass through a filter with a loss of less than 3 dB relative to the nominal gain of the filter.

23.1.2 Stop Band

The *stop band* is a band of frequencies that pass through a filter with a loss of greater than 3 dB relative to the nominal gain of the filter.

23.1.3 Cutoff Frequency

A *cutoff frequency* is the frequency at which the gain first falls to 3 dB below the nominal gain of the filter, as you move out of the pass band.

23.1.4 Corner Frequency

A *corner frequency* is the frequency at which the rate of change of a response makes a noticeable change. In the case of a low-pass or high-pass filter, this is the same as a the cutoff frequency, but other filters such as shelving filters may have additional corner frequencies.

23.1.5 Bandwidth

The *bandwidth* is the difference between the upper and lower cutoff frequencies on either side of the pass band.

23.1.6 Transition Band

The *transition band* is the range of frequencies over which the gain the filter falls from its level at the cutoff frequency to the nominal attenuation level in the stop band.

23.1.7 Center Frequency

The *center frequency* of a band of frequencies is defined as the geometric mean of the lowest and highest frequencies of the band.

$$f_m = \sqrt{f_1 \times f_2} \tag{23-1}$$

where,

f_1 is the cutoff frequency of the high-pass filter,

f_2 is the cutoff frequency of the low-pass filter.

23.1.7.1 Geometric Symmetry

A response showing mirror image symmetry about the center frequency when plotted on a log scale is said to have *geometric symmetry.* This is the natural response of many electrical circuits as the response function tends to contain multiplicative terms.

23.1.7.2 Arithmetic Symmetry

A response showing mirror image symmetry about the center frequency when plotted on a linear scale is said to have *arithmetic symmetry*. A bandpass filter with a constant envelope delay will have arithmetic symmetry in both phase and amplitude. The center frequency in this case will be given by the arithmetic mean

$$f_c = \frac{f_1 + f_2}{2} \tag{23-2}$$

23.1.8 Order

The *order* of a filter is determined by the number of *reactive* elements in the circuit. These can either be *inductive* or *capacitive* and generally only include those added for purposes of the frequency response within the audio band and not for stability or RF suppression. If all of the elements act as either low pass or high pass, the roll off in the stop band will approach 6 dB per octave per order. A fourth-order low pass will have a roll-off of 24 dB per octave above the cutoff frequency, but a fourth-order band pass will have 12 dB per octave on either side of the center frequency.

23.1.9 Phase Angle

The *phase angle* at a particular frequency is a measure of the relative time for a particular frequency to pass through a system from input to output. Phase angle is a relative measure and is usually expressed in degrees where 360° represents one wavelength. In most formulas, phase is used in terms of radians where 2π represents one wavelength. The instantaneous phase of a sinusoidal signal is given by

$$\begin{aligned}\alpha &= \omega t \\ &= 2\pi \times ft\end{aligned} \tag{23-3}$$

23.1.10 Phase Delay

The *phase delay* of a system at a given frequency is the equivalent time offset that would induce the same phase offset as measured on a sinusoid of the same frequency.

$$\begin{aligned}\tau_p &= \frac{\alpha}{\omega} \\ &= \frac{\alpha}{2\pi f}\end{aligned} \tag{23-4}$$

23.1.11 Group Delay

A filter can exhibit a *group delay* over a group of frequencies covering a section of the audio spectrum if those frequencies are all subject to the same time delay. The group delay is given as the first derivative of the phase with respect to frequency

$$\tau_g = \frac{d}{d\omega}\phi(\omega)\,. \tag{23-5}$$

The threshold of perceptibility for group delay has been shown to be between 1 to 3 ms over the 500 Hz to 4 kHz range of the audio spectrum.[1]

23.1.12 Transient Response

The *transient response* of a filter is the time response to an input stimulus. *Impulse* and *step* inputs are common stimuli for this measurement. Narrow bandwidth filters, when subjected to rapidly changing input, ring because it takes a certain amount of time for the energy in the network to change upon application or removal of the signal. Ringing can most clearly be seen as a damped tail on a signal after it has been removed, Fig. 23-2.

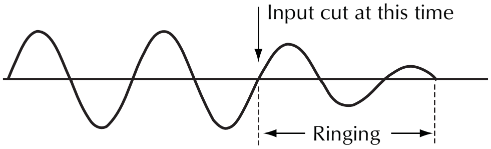


Figure 23-2. Ringing of a filter after the removal of a signal.

23.1.13 Minimum Phase

A *minimum phase* system is one for which the phase shift at each frequency can be uniquely determined from the magnitude response using the *Hilbert transform*. A filter with more than one path from input to output, in which the different branches have a different *group delay,* will be a *linear time invariant* (LTI) system but may be nonminimum phase.

23.2 Passive Filters

Passive filters do not have any amplification components in the circuit and as such cannot put out more energy than is put in. A passive filter can never have a boost in the energy response, although with some resonant circuits, instantaneous voltages may be higher than the input voltage. In this case the output impedance will rise, preventing any significant current from being driven. To build a passive filter with boost, we must construct a filter that cuts all other frequencies and then use a separate amplifier to increase the overall gain.

23.2.1 First-Order L and C Networks

Inductor- and capacitor-based filter networks may be analyzed in terms of their impedances by reducing the circuit to its resistance and reactance components.

The impedance may be represented as a single complex number where the real part is the resistance and the imaginary part is the reactance. The imaginary part of a complex number is given by the magnitude multiplied by the square root of negative one. The mathematical notation for this number is *i* but in engineering, *j* is commonly used to avoid confusion in expressions involving current.

$$Z = R + jX \tag{23-6}$$

Analyzing a network in term of complex impedance allows the calculation of both magnitude and phase at any frequency according to

$$\theta = \tan^{-1}\left(\frac{imaginary}{real}\right) \tag{23-7}$$

$$A = \sqrt{imaginary^2 + real^2} \tag{23-8}$$

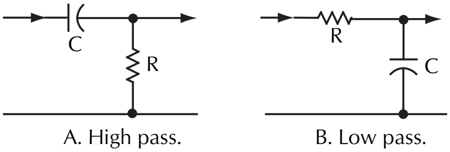
where,
θ is the phase angle of the complex number,
A is the magnitude of the complex number.

23.2.1.1 Capacitive Networks

The capacitor has impedance that approaches a short circuit at high frequency and an open circuit at low frequency. The reactance of a capacitor is given by:

$$X_C = \frac{1}{2\pi fC} \tag{23-9}$$

where;
X_c is the capacitive reactance,
f is the frequency in hertz,
C is the capacitance in farads.



A. High pass. B. Low pass.

Figure 23-3. Simple filter networks using only a capacitor and a resistor.

If a capacitor is connected in series with the signal path as in Fig. 23-3A, the capacitor and the resistor form a *potential divider*. Low frequencies will be attenuated as the impedance of the capacitor increases at lower frequencies.

$$V_{out} = V_{in}\frac{R}{(R + Z_c)} \tag{23-10}$$

The cutoff frequency of this filter is at the frequency where $R = |Z_C|$, so substituting into Eq. 23-8, we find

$$f = \frac{1}{2\pi RC} \tag{23-11}$$

Using the complex analysis in Eq. 23-8, we can determine the phase at this frequency.

$$\begin{aligned} V_{out} &= \frac{V_{in} \times R}{R + jR} \\ &= \frac{1}{1 + j} \\ &= \frac{1 - j}{2} \end{aligned}$$

So according to Eqs. 23-7 and 23-8, the magnitude is 0.707 or −3 dB and the phase angle is −45°.

If a capacitor is connected in parallel with the signal path as in Fig. 23-3B, the capacitor and the resistor form a *potential divider*. High frequencies will be attenuated as the impedance of the capacitor reduces at higher frequencies.

$$V_{out} = V_{in}\frac{Z_c}{(R + Z_C)} \tag{23-12}$$

The cutoff frequency of this filter is at the frequency where $R = |Z_C|$, so substituting into Eq. 23-10, we find that again

$$f = \frac{1}{2\pi RC} \tag{23-13}$$

Using the complex analysis in Eq. 23-10, we can determine the phase at this frequency.

$$\begin{aligned} V_{out} &= \frac{V_{in} \times jR}{R + jR} \\ &= \frac{j}{1 + j} \\ &= \frac{1 + j}{2} \end{aligned}$$

So according to Eqs. 23-7 and 23-8, the magnitude is 0.707 or −3 dB and the phase angle is +45°.

23.2.1.2 Inductive Networks

The inductor has impedance that approaches an open circuit at high frequency and a short circuit at low frequency. The reactance of an inductor is given by

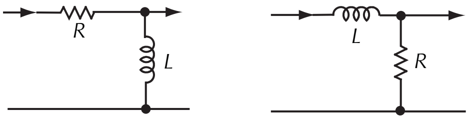
$$X_L = 2\pi f L \tag{23-14}$$

where,
X_L is the inductive reactance in ohms,
f is the frequency in hertz,
L is the inductance in henrys.

Inductors are prone to parasitic resistances especially for large inductances where a long coil is wound. In a large value inductor, the parasitic resistance is reduced by using heavier gauge wire, which causes the size to grow rapidly as the inductance becomes larger. The full expression for the impedance of an inductor is

$$Z_L = R_L + j2\pi f L \tag{23-15}$$

where,
Z_L is the impedance of the inductor,
R_L is the dc resistance of the inductor.



A. High pass. B. Low pass.

Figure 23-4. Filters using only an inductor and a resistor.

If an inductor is connected in series with the signal path as in Fig. 23-4B, the inductor and the resistor form a *potential divider*. High frequencies will be attenuated as the impedance of the inductor increases at higher frequencies.

$$V_{out} = V_{in}\frac{Z_c}{(R + Z_L)} \tag{23-16}$$

The cutoff frequency of this filter is at the frequency where $R = |Z_L|$, so substituting into Eq. 23-14, we find that

$$f = \frac{L}{2\pi R} \tag{23-17}$$

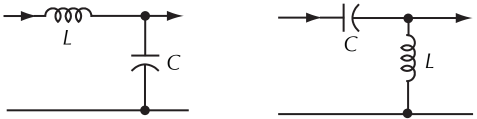
Using the complex analysis in Eq. 23-10 and ignoring the parasitic resistance, we can determine the phase at this frequency.

$$\begin{aligned} V_{out} &= \frac{V_{in} \times R}{R + jR} \\ &= \frac{j}{1 + j} \\ &= \frac{1 - j}{2} \end{aligned}$$

So according to Eqs. 23-7 and 23-8 magnitude is 0.707 or –3 dB and the phase angle is –45°. Note that this is the opposite phase angle to the capacitor-based low-pass filter.

23.2.2 Second-Order L-Type Networks

An L-type filter consists of an inductor in series with a capacitor, with the outputs across one or more of the components. Since there are two reactive elements in the circuit, it forms a *second-order* filter with a roll-off of 12 dB per octave. There are two configurations of this network.



A. High-frequency attenuated, inverted L-type filter. B. Low-frequency attenuated, inverted L-type filter.

Figure 23-5. Two configurations of L-type filters.

The insertion loss for the low-pass configuration as shown in Fig. 23-5A is given by

$$IL_{dB} = 10\log\left[1 + \left(\frac{f}{f_c}\right)^2\right]. \tag{23-18}$$

The insertion loss for the high-pass configuration as shown in Fig. 23-5B is given by

$$IL_{dB} = 10\log\left[1 + \left(\frac{f_a}{f}\right)^2\right] \tag{23-19}$$

where,
f_c is the frequency of a 3 dB insertion loss,
f is any frequency,
IL_{dB} is the insertion loss in decibels.

These configurations are commonly used in basic loudspeaker crossover networks as in Fig. 23-6. Both a high-pass and a low-pass response may be derived from the same circuit. The L-type filter in this application presents constant impedance to the input port. The

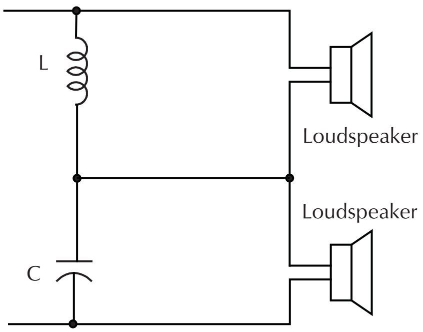


Figure 23-6. Passive crossover using an L-type filter.

impedance of the inductor Eq. 23-12 and the capacitor Eq. 23-7 vary with frequency and are chosen such that, at the crossover frequency, their impedances equal the characteristic impedance, Z_0. Each port is in parallel with a load, which for simplicity of analysis we will consider to be constant and of value Z_0:

$$Z_L = \frac{1}{\left(\frac{1}{2\pi f_x L} + \frac{1}{Z_0}\right)} \tag{23-20}$$

$$Z_C = \frac{1}{\left(2\pi f_x C + \frac{1}{Z_0}\right)} \tag{23-21}$$

where,

Z_0 is the circuit impedance,

F_x is the crossover frequency.

The total impedance at the input is

$$Z_{in} = Z_C + Z_L\,. \tag{23-22}$$

When the frequency is very much lower than the crossover frequency, the value of Z_L becomes $2\pi f_x L$, which is very small. At the same time, the value of Z_C becomes Z_0 as $2\pi f_x C$ becomes smaller. The total impedance becomes Z_0.

At the crossover frequency, the inductor and capacitor impedances equal Z_0, so the total circuit impedance also equals Z_0.

When the frequency is very much higher than the crossover frequency, the value of Z_L becomes Z_0 as $2\pi f_x L$ becomes larger. At the same time, the value of Z_C becomes $1/(2\pi f_x C)$, which is small. The total impedance becomes Z_0.

23.2.3 T and π Networks

T and *π* networks are classes of *constant-k* filters. They are formed by combing L-type filters with one leg being in common. The line impedance Z_0 is a critical parameter in the design of these filters. The impedance presented by a T network to the input and output transmission lines is symmetrical and is designated Z_T. This impedance is equal to the line impedance in the passband and progressively decreases in the stop band. The impedance presented by a π network to the transmission lines is also symmetrical and is designated Z_P. This impedance is equal to the line impedance in the passband and progressively increases in the stop band.

The full T and π networks have twice the attenuation of the L-type half sections.

23.2.3.1 Low Pass

A T-type *low-pass* filter has two inductances: L_1 in series with the line and a capacitance C_2 in parallel. As frequency increases, the inductive reactance increases, presenting an increasing opposition to transmission. As frequency increases, capacitive reactance reduces, so the parallel capacitor becomes more effective at shunting the signal to ground. The design equations for the component values are

$$C_2 = \frac{1}{2\pi f_c Z_0} \tag{23-23}$$

$$L_1 = \frac{Z_0}{2\pi f_c} \tag{23-24}$$

where,

f_c is the cutoff frequency,

Z_0 is the line impedance.

These equations are the same as for the L-type network. In the T network, the actual value of the capacitor is $2C_2$, where the capacitors from two low-pass L-type networks are combined in parallel. In the π network, the actual value of the indictor is $2L_1$, where the inductors from the L-type network are combined in series.

23.2.3.2 High Pass

The basic designs of constant-k high-pass filters are shown in Fig. 23-8. The positions of the inductors and capacitors are opposite to those in the low-pass case. The design equations are

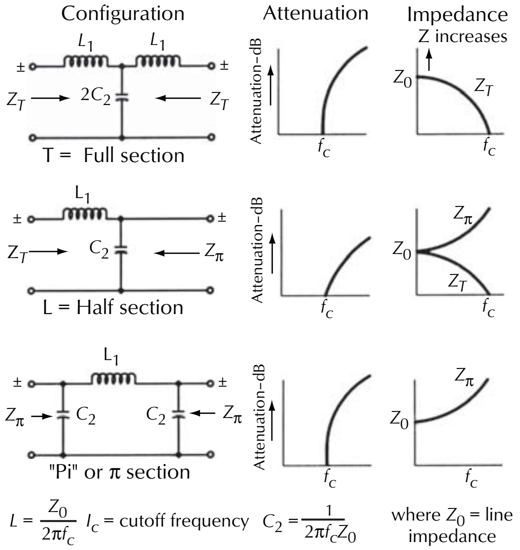

Configuration
Attenuation
Impedance
Z increases
L_1
L_1
±
±
Z_T
$2C_2$
Z_T
T = Full section
Attenuation–dB
f_c
Z_0
Z_T
f_c
L_1
Z_T
C_2
Z_π
L = Half section
Z_π
Z_0
Z_T
L_1
Z_π
C_2
C_2
Z_π
"Pi" or π section
Z_0
Z_π
$L = \frac{Z_0}{2\pi f_c}$
l_c = cutoff frequency
$C_2 = \frac{1}{2\pi f_c Z_0}$
where Z_0 = line impedance

Figure 23-7. Configuration and characteristics of low-pass filters.

$$C_1 = \frac{1}{2\pi f_c Z_0} \tag{23-25}$$

$$L_2 = \frac{Z_0}{2\pi f_c} \tag{23-26}$$

where,

f_c is the cutoff frequency,

Z_0 is the line impedance.

23.2.3.3 Parallel Resonant Elements

A *parallel resonant* circuit element has impedance that is at a maximum at the resonant frequency (Fig. 23-9). The impedance of the element is given by

$$Z = \frac{X_L \times X_C}{X_L + X_C} \tag{23-27}$$

where,

Z is the impedance,

X_L is the reactance of the inductor,

X_C is the reactance of the capacitor.

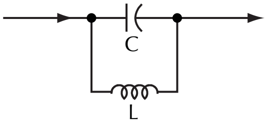


Figure 23-9. Parallel resonant circuit.

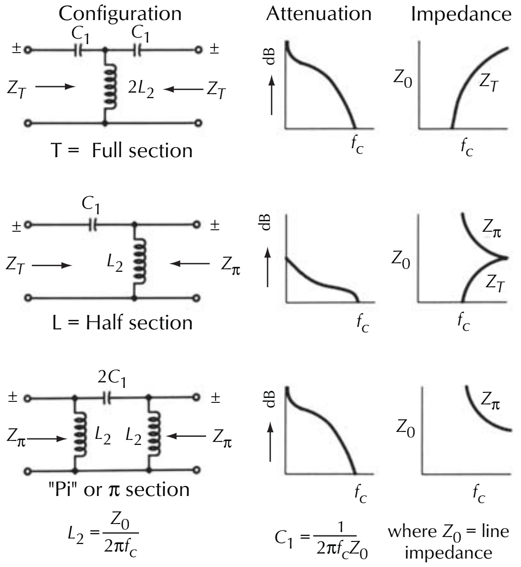

Configuration
Attenuation
Impedance
C_1
C_1
±
±
Z_T
$2L_2$
Z_T
T = Full section
dB
f_c
Z_0
Z_T
f_c
C_1
Z_T
L_2
Z_π
L = Half section
Z_π
Z_0
Z_T
$2C_1$
Z_π
L_2
L_2
Z_π
"Pi" or π section
Z_0
Z_π
$L_2 = \frac{Z_0}{2\pi f_c}$
$C_1 = \frac{1}{2\pi f_c Z_0}$
where Z_0 = line impedance

Figure 23-8. Configuration and characteristics of high-pass filters.

At very low frequencies, the reactance of the inductor approaches a short circuit, reducing the overall impedance. At high frequencies, the reactance of the capacitor approaches a short circuit, reducing the overall impedance.

23.2.3.4 Series Resonant Elements

A *series resonant* circuit element has impedance that is at a minimum at the resonant frequency. The impedance of the element is given by

$$Z = X_L + X_C \tag{23-28}$$

where,

Z is the impedance,

X_L is the reactance of the inductor,

X_C is the reactance of the capacitor.

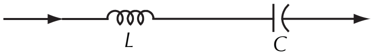


Figure 23-10. Series-resonant circuit.

At very low frequencies, the reactance of the capacitor approaches an open circuit, increasing the overall impedance. At high frequencies, the reactance of the inductor approaches an open circuit, increasing the overall impedance.

23.2.3.5 Bandpass

The impedance characteristics of the *series* and *parallel* resonant elements can be used to form a bandpass filter as in Fig. 23-11. The frequencies f_1 and f_2 are the cutoff frequencies of the pass band. The design equations for the component values are

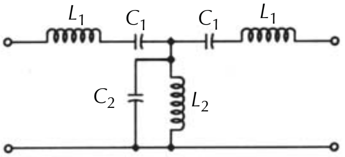
$$L_1 = \frac{Z_0}{2\pi(f_2 - f_1)} \tag{23-29}$$

$$L_2 = \frac{(f_2 - f_1)Z_0}{2\pi f_1 f_2} \tag{23-30}$$

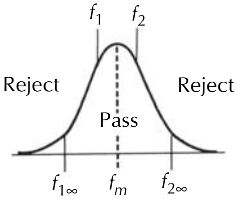
$$C_1 = \frac{f_2 - f_1}{2\pi f_1 f_2 Z_0} \tag{23-31}$$

$$C_2 = \frac{1}{2\pi(f_2 - f_1)Z_0} \tag{23-32}$$

where,
f_1 is the lower cutoff frequency,
f_2 is the upper cutoff frequency,
Z_0 is the line impedance.



A. Configuration.



B. Transmission characteristics.

Figure 23-11. T network bandpass filter.

23.2.3.6 Band Reject

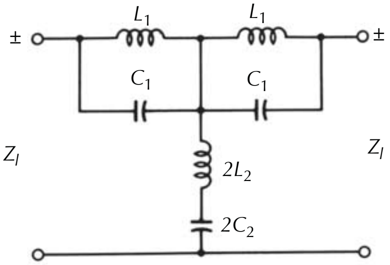
The configuration for a *band reject* filter using series and parallel resonant elements is shown in Fig. 23-12. The configuration is the reverse of the bandpass T-network filter. In this case the frequencies f_1 and f_2 are at the edge of the reject band. The design equation for the component values are

$$L_1 = \frac{(f_2 - f_1)Z_0}{2\pi f_2 f_1} \tag{23-33}$$

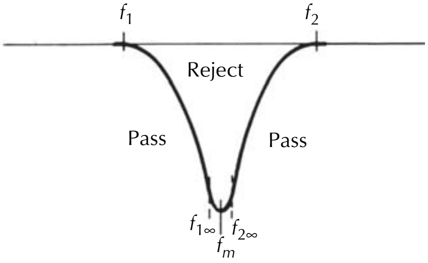
$$C_1 = \frac{1}{2\pi(f_2 - f_1)Z_0} \tag{23-34}$$

$$L_2 = \frac{Z_0}{2\pi(f_2 - f_1)} \tag{23-35}$$

$$C_2 = \frac{f_2 - f_1}{2\pi f_1 f_2 Z_0} \tag{23-36}$$



A. Configuration.



B. Transmission characteristics.

Figure 23-12. T network band-reject filter.

23.2.3.7 Ladder Networks

Passive filters of arbitrary length may be constructed by adding RC, RL, or LC L-type sections into a network of arbitrary length called a *Cauer network*. The interaction between the various stages in this topography starts to become important as the impedance of one section loads the next.

23.2.4 Filter Design

As the number of components in a filter increases, the number of possible transfer functions also increases. Increasing the order of a filter by adding more of the same sections will not necessarily produce the optimum results. Consider chaining two low-pass filters with a cutoff frequency of f_c. The attenuation at the cutoff is 3 dB, so with two sections in series, the attenuation at f_c

is 6 dB. This means that the 3 dB cutoff point has moved somewhat lower.

We can analyze a filter in the *Laplace domain* in terms of input signals of the form e^{st} with s defined as

$$s = \sigma + j\omega \tag{23-37}$$

where,

σ is a value for exponential decay,

ω is $2\pi f$, f being the frequency.

This gives us a transfer function that can be expressed as polynomial functions in s. A first order low-pass filter is of the form

$$h_1(s) = \frac{1}{(s+p_0)} \tag{23-38}$$

The value of p_0 defines the cutoff frequency. Adding more sections in series progressively multiplies more terms,

$$h_1(s) = \frac{1}{(s+p_0)(s+p_1)} \tag{23-39}$$

For a *normalized* version of Eq. 23-37, p_0 is set to be one, and all other values in the sequence of p_n can be defined according to a formula. The exact formula used depends on the most important characteristic of the filter you are designing.

23.2.4.1 Butterworth

The *Butterworth* filter is maximally flat and has the most linear phase response in the pass band but has the slowest transition from pass band to stop band for a given order. The polynomial transfer function in the form of Eq. 23-37 can be constructed using a formula.

$$B_n(s) = \prod_{k=1}^{\frac{n}{2}} \left[s^2 + 2\cos\left(\frac{2k+n-1}{2n}\pi\right)s + 1 \right] \tag{23-40}$$

Eq. 23-40 gives the polynomials for an even order of filter. To calculate the polynomial for an odd order, add a term $(s + 1)$, and then apply the formula with $n = n–1$. Table 23-1 gives the calculated values for the Butterworth polynomials up to fifth order.

Table 23-1. Butterworth Polynomials

Order	Polynomial
1	$(s + 1)$
2	$(s^2 + 1.414s + 1)$
3	$(s^2 + 1)(s^2 + s + 1)$
4	$(s^2 + 0.765 + 1)\ (s^2 + 1.848s + 1)$
5	$(s + 1)(s^2 + 0.618s + 1)(s^2 + 1.618s + 1)$

23.2.4.2 Linkwitz-Riley

The *Linkwitz-Riley* filter[2] is used in audio crossovers. It is formed by cascading two Butterworth filters so that the cutoff at the crossover frequency is –6 dB. This means that summing the low-pass and high-pass responses will have a gain of 0 dB at crossover and all other points.

23.2.4.3 Chebyshev I and II

Chebyshev filters have a steeper roll-off than the Butterworth filters but at the expense of a ripple in the response. There are two forms of the Chebyshev filter. Type I has a ripple in the pass band and maximum attenuation in the stop band. Type II is the reverse, with a flat pass band and a ripple in the stop band that limits the average attenuation.

The filter's transfer function is defined in terms of a ripple factor ε

$$H(\omega) = \frac{1}{\sqrt{1 + \varepsilon^2 C_n^{\ 2} \frac{\omega}{\omega_0}}} \tag{23-41}$$

where,

C_n is the polynomial for the order n, as given in Table 23-2.

The magnitude of the ripple in decibels is

$$ripple_{dB} = 20\log\left(\frac{1}{\sqrt{1+\varepsilon^2}}\right) \text{ dB} . \tag{23-42}$$

Table 23-2. Chebyshev Polynomials

Order	Type I	Type II
1	s	$2s$
2	$2s^2 – 1$	$4s^2 – 1s$
3	$4s^3 – 3s$	$8s^3 – 4s$
4	$8s^4 – 8s^2 + 1$	$16s^4 – 12s^2 + 1$
5	$16s^5 – 20s^3 + 5s$	$32s^5 – 32s^3 + 6s$

23.2.4.4 Elliptical

The *elliptical* filter has a ripple in both the pass band and the stop band, with the shortest possible transition band for the order of the filter with a given ripple. The ripple in the pass band and the stop band are independently controllable. This is a generalized form of the *Butterworth* and *Chebyshev* filters. If the pass band and stop band ripple is set to zero, we have a Butterworth filter. If the pass band has a ripple and the stop band does not, we have a Chebyshev Type I. If the stop band has a ripple and the pass band does not, we have a Chebyshev Type II.

The transfer function is the same form as Eq. 23-39 with a different polynomial

$$H_n(\omega) = \frac{1}{\sqrt{1+\varepsilon^2 E_{n,\zeta}{}^2\left(\frac{\omega}{\omega_0}\right)}} \qquad (23\text{-}43)$$

where,

$E_{(n,\xi)}$ is the elliptical polynomial for the order n and selectivity factor ξ.

23.2.4.5 Normalizing

Normalizing is the process of adjusting the values of filter components to a convenient frequency and impedance. For analysis, the frequency is usually normalized to 1 rad s^{-1} and the impedance to 1 Ω. For designing practical audio circuits the filter is normalized to 1 kHz and 10 kΩ.

23.2.4.6 Scaling

Scaling is the design process of changing the normalized frequency or impedance values for a filter by varying resistor and capacitor values. Frequency can be changed relative to the normalized frequency by either changing all of the resistor values or all of the capacitor values by the ratio ρ of the desired frequency to the normalized frequency. From Eq. 23-11, frequency varies inversely with the product of the capacitor and resistor value.

$$\rho = \frac{f_{norm}}{f_1} \qquad (23\text{-}44)$$

where,

ρ is the scaling factor,

f_1 is the new frequency.

By multiplying all of the resistor values by a factor, and dividing all of the capacitor values by that same factor, we can change the normalized impedance of the network without changing the RC product, thus keeping the frequency unchanged.

$$\rho = \frac{Z_1}{Z_{norm}} \qquad (23\text{-}45)$$

where,

ρ is the scaling factor,

Z_1 is the impedance.

23.2.5 Q and Damping Factor

A *damping factor*, d, or its reciprocal, Q, appears in the design equation of some filters. The circuit behaves differently depending on the value of d.

When d is 2, the damping is equivalent to the isolated resistance-capacitance filters.

When d is 1.41 (square root of 2), the filter is *critically damped* and gives maximum flatness without overshoot.

As d decreases between 1.414 and 0, the overshoot peak increases in level with its being 1 dB at $d = 1.059$, 3 dB at $d = 0.776$.

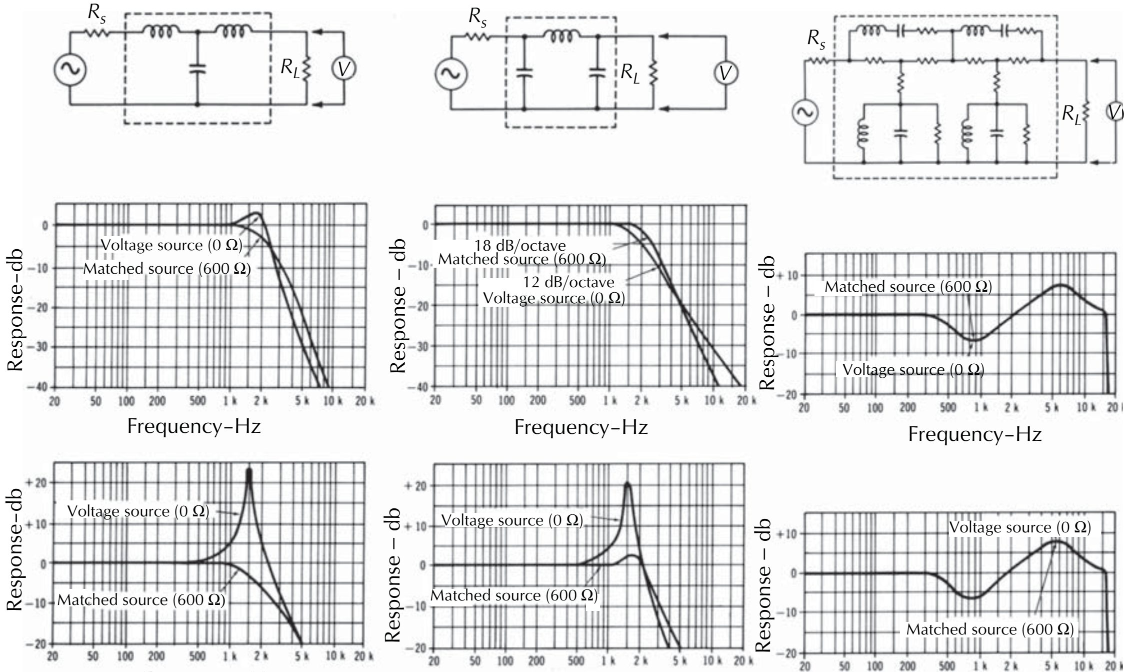
When d is 0, the peak becomes so large that the filter becomes unstable, and if gain is applied it can become an oscillator.

23.2.6 Impedance Matching

Source and load impedance have an effect on a passive filter's response. They can change the cutoff frequency, attenuation rate, or Q of the filter. Fig. 23-13 shows the effects of improper source and load impedance on three different passive filters. The peaks in the response before the cutoff frequency lead to a ringing in the filter, making it potentially unstable at these frequencies. The bridged T filter is not affected by the impedance mismatch because of the resistors in the filter; however, these resistors create an insertion loss.

23.3 Active Filters

Any passive filter may be turned into an active filter by using amplification at the input and output to provide the option of gain, Fig. 23-14. This also provides important buffering, giving the circuit a high-input impedance and low-output impedance, guarding the circuit against external impedance mismatches. This allows active



A. T network. B. π network. C. Bridged-T network.

Figure 23-13. Effects of termination impedance on three types of filter sections.

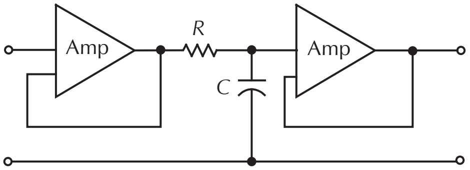


Figure 23-14. Simple buffering of an active filter.

filter sections to be connected together without concerns for mutual interference.

More advanced active filters use filter components in the feedback loop of a gain stage to add functionality with fewer components. Active filters have advantages over passive filters in that they can be made much smaller, especially for low-frequency filters that would otherwise use bulky inductors. The removal of inductors also makes active filters less prone to low-frequency hum interference. The disadvantages of active filters are that they are more complex, having more components to fail; require a power supply; and have a dynamic range limited at the top by the power supply and at the bottom by high-frequency self-noise in the amplifiers.

23.3.1 Filter Topologies

23.3.1.1 Sallen-Key

Sallen-Key filters are second-order high-pass or low-pass sections exhibiting a 12 dB per octave cutoff slope in the stop band. *Equal component value* filters are the easiest to design, with the frequency-determining resistors being of equal value and the frequency-determining capacitors being of equal value. They have the advantage of being able to high pass or low pass simply by interchanging their positions.

In the second-order low pass of Fig. 23-15, frequency is changed by scaling the values of R and C in the input network in accordance with Eq. 23-42. To keep the offset at a minimum, it is best to have R_0 equal to the input impedance of $2R$. Damping factor, d, is controlled by the ratio of R_f and R_0 such that

$$R_f = (2-d)R_0 \tag{23-46}$$

The gain of the circuit is fixed at

$$\begin{aligned} gain &= 1+\frac{R_f}{R_0} \\ &= \frac{1+(2-d)R_0}{R_0} \qquad (23\text{-}47) \\ &= (3-d) \end{aligned}$$

where

d is the damping factor,

R_f is the op-amp feedback resistance,

R_0 is the resistance between ground and the inverting input.

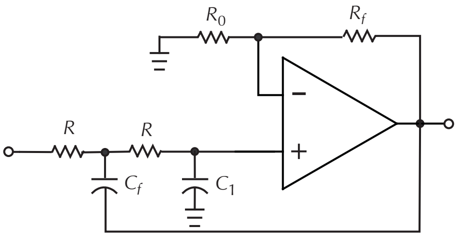


Figure 23-15. Sallen-Key low-pass filter.

The second-order high-pass filter of Fig. 23-16 is constructed by reversing the locations of *R* and *C* in Fig. 23-15. The gain and damping factor follow the same equations as for the low pass.

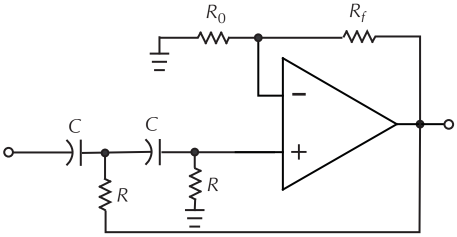


Figure 23-16. Sallen-Key high-pass filter.

A *unity gain* Sallen-Key filter can also be made. To independently control frequency and damping, the ratio of the capacitors must be changed such that in the low pass

$$C_f = \left(\frac{4}{d5}\right)C_1 . \qquad (23\text{-}48)$$

The cutoff frequency is still determined by the product of *R* and *C*, so it can be adjusted with the value of *R* or by scaling C_f and C_1 together.

Fig. 23-17 is a Sallen-Key filter implemented as a bipolar junction transistor circuit.

23.3.1.2 State Variable

The *state variable* filter consists of two low-pass filters and a summing stage. High-pass, bandpass, and low-pass outputs are all available from the circuit. The operation relies on both the magnitude and phase characteristics of the low-pass sections to generate the outputs.

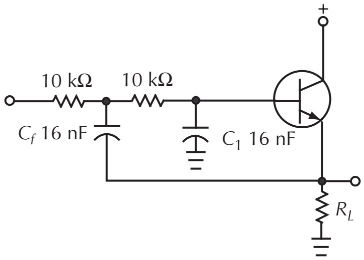


Figure 23-17. Sallen-Key filter implemented as a Bipolat Junction Transistor (BJT) circuit.

At high frequency, the low-pass sections attenuate the signal so that the feedback signal is small, leaving the unaffected signal at the high-pass output. As the input frequency approaches the center frequency, the levels at both the bandpass and low-pass outputs begin to increase. This leads first to an increase in positive feedback from the bandpass section giving a damping dependent overshoot. When the input frequency is below the center frequency, the net phase shift of both low-pass sections is 180 degrees, leading to negative feedback and an attenuation of the high-pass output.

The cutoff frequency of the filter in Fig. 23-18 can be changed as in the preceding circuits by varying R_1 and R_2 or C_1 and C_2 while keeping other values identical. The damping factor is varied by changing the band-pass feedback gain, controlled by the ratio of R_3 and R_4.

$$d = \frac{R_4}{R_3} \qquad (23\text{-}49)$$

The overall gain is controlled by R_{12}. If R_1, R_2, and R_{12} are equal, the gain is one.

$$gain = \frac{R_1}{R_{12}} \qquad (23\text{-}50)$$

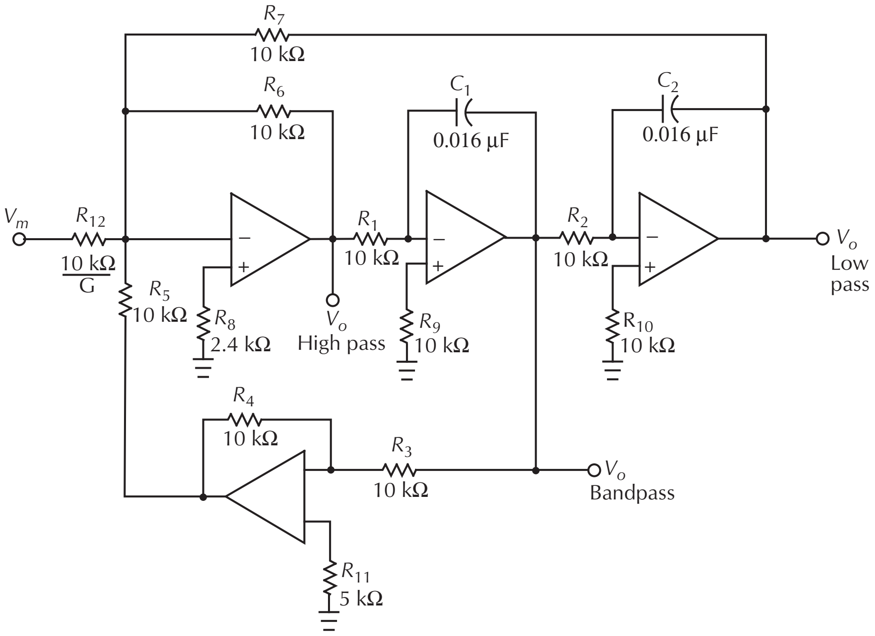


Figure 23-18. Second-order state variable filter.

The values of R_8, R_9, R_{10}, and R_{11} are not critical and should be chosen for minimum dc offset at each op-amp stage.

23.3.1.3 All-Pass Filter

The circuit shown in Fig. 23-19 is an all-pass amplifier with unity gain at all frequencies and having a phase shift proportional to frequency according to

$$\theta = 2\tan^{-1}\left(\frac{f_0}{f}\right) \tag{23-51}$$

where,

θ is the phase shift from input to output,

f_0 is $1/(2\pi RC)$.

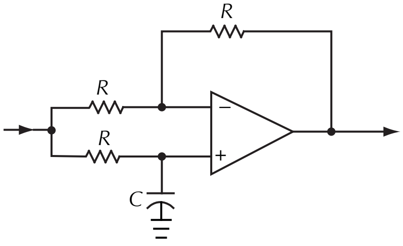


Figure 23-19. All pass unity gain amplifier.

The phase shift is approximately proportional to the frequency over a range of frequencies below and above f_0. These circuits can be cascaded to induce more phase-shift over the same frequency range or each designed with a different f_0 to extend the range over which phase is proportional to frequency.

$$\begin{aligned}\theta &= \omega t \\ &= 2\pi f t\end{aligned} \tag{23-52}$$

Since phase is proportional to frequency, and from Eq. 23-50 phase is an expression of time, these circuits may be used to introduce a small amount of delay.

23.3.2 Pole-Zero Analysis

A pole-zero plot, Fig. 23-20 is graphical way of representing the complex transfer function of a filter. The pole-zero plot describes a surface that has peaks of infinite magnitude that stretch the surface upward and zeros that do the same downward. The height of the surface along the ω axis, where $\sigma = 0$, is the normal magnitude response.

If the expression for the function is reduced to a factored form in the *s-plane* where s is the Laplace domain variable Eq. 23-35, then the transfer function of a system can be represented as

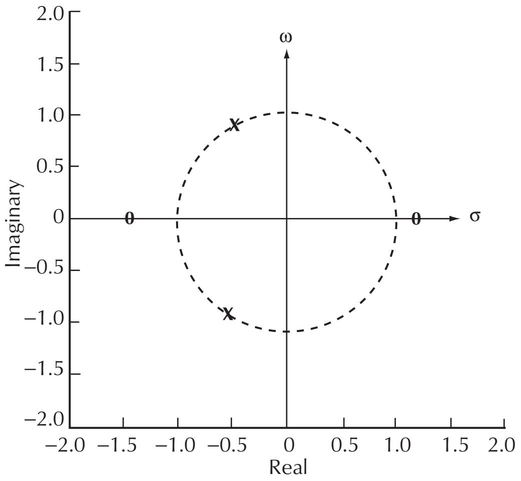


Figure 23-20. Pole-zero plot with poles represented as X and zeros as 0.

$$H(s) = \frac{P(s)}{Q(s)} \tag{23-53}$$

where,

$P(s)$ and $Q(s)$ are polynomials expressible in the form
$P(s) = (s - p_1)(s - p_2)\ldots(s - p_n)$,
$Q(s) = (s - q_1)(s - q_2)\ldots(s - q_n)$.

23.3.2.1 Zeros

The *zeros* of the function are the values of s at which $P(s)$ is zero and consequently $H(s)$ is zero. These occur at values p_1, p_2, and so on, and represent frequencies at which the transfer function exhibits maximum attenuation.

23.3.2.2 Poles

The *poles* of the function are the values of s at which $Q(s)$ is zero and consequently $H(s)$ is infinite. These occur at values q_1, q_2, and so on, and represent frequencies at which the transfer function exhibits maximum gain.

23.3.2.3 Stability

A pole or a zero in the right-hand side of the s-plane means that for that value of s, σ is greater than zero. In the time domain representation, the signal is given as

$$f(t) = \int_0^{\infty} e^{st} F(s) ds \tag{23-54}$$

The term e^{st} may be expanded to $e^{\sigma t} \times e^{j\omega t}$. If the value of σ is greater than zero, the expression represents an exponentially increasing factor, meaning that the filter is unstable. This situation cannot arise in a passive filter so they are inherently stable.

23.4 Switched Capacitor Filters

Any active filter based on resistive and capacitive components may be reconfigured as a *switched capacitor filter*. The resistive elements are replaced by an equivalent switched capacitive element. The advantages of using switched capacitors in place of resistors is that they are easier to implement in silicon, since capacitors take up less space than resistors, and tolerances of capacitor-to-capacitor ratios can be more easily controlled the resistor-capacitor products.

The circuit shown in Fig. 23-21 transfers charge, and therefore current, between the two voltage sources under control of the switch. The charge ΔQ transferred every switch period of length t_s may be expressed in terms of current Eq. 23-53 or voltage Eq. 23-54.

$$\begin{aligned} \Delta Q &= I t_s \\ &= \frac{I}{f_s} \end{aligned} \tag{23-55}$$

$$\Delta Q = C(v_1 - v_2) \tag{23-56}$$

Combining these two equations we can find the equivalent resistance.

$$\frac{I}{f_s} = C(v_1 - v_2)$$

$$\begin{aligned} R &= \frac{(v_1 - v_2)}{I} \\ &= \frac{I}{C f_s} \end{aligned} \tag{23-57}$$

The equivalent resistor value in Eq. 23-55 has a fixed capacitive term and a frequency term. Its value may be controlled by varying the switching frequency. This makes switched capacitor filters ideal for filters that need to be tuned.

23.5 Digital Filters

Filters may be implemented using entirely mathematical means from their transfer function representations in the time domain. The time and frequency domains are

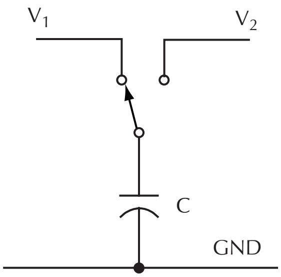


Figure 23-21. Switched capacitor equivalent of a resistor.

related by the *Fourier transform*. Digital filters make use of extensively recursive algorithms involving multiplications and additions, for which digital signal processors (DSPs) are optimized. The precision of the sampled data in magnitude and time is an important factor, not only at the input and output but all through the calculations.

23.5.1 FIR Filters

A *finite impulse response* (FIR) filter performs the convolution in the time domain of the input signal and the impulse response of the filter. While FIR filters are simple in concept and easy to design, they can end up using large amounts of processing power relative to other designs. Hundreds of multiplications per sample are often needed. They are, however, inherently stable as there are no feedback loops that can get out of control when finite precision arithmetic is used. They can also be designed to have *linear phase*, preserving wave shape and having a constant time delay for all frequency components.

The FIR filter structure is shown in Fig. 23-22. Each Z^{-1} is a delay that represents one unit of time equivalent to the sample period of the system. The notation derives from the Z domain transform, which is a way of expressing transfer functions in a discrete time form. The recursive nature of the algorithm is apparent, with the multiply and add sections being repeated for every sample in the stored impulse response.

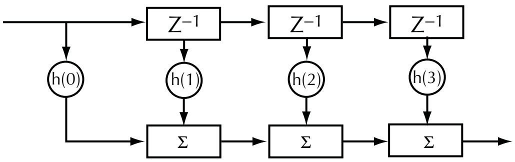


Figure 23-22. Block diagram of an FIR filter.

The FIR filter treats each incoming sample as an input impulse stimulus and generates an output that is a truncated copy of the impulse response scaled by the magnitude of that sample. The summing of the results from each successive sample by *superposition* generates the full output signal. The result for each output sample in a filter with M coefficients is

$$y(n) = \sum_{m=0}^{M} x(n-m) \times h(m) \tag{23-58}$$

where,
n is the sample number,
$x(n)$ is the nth input sample value,
$h(m)$ is the mth filter coefficient value,
$y(n)$ is the nth output sample value.

This requires the storage of $M - 1$ previous input samples and is executed in M multiply and add operations per sample.

23.5.1.1 FIR Coefficients

The coefficient values for an FIR filter are generally computed in advance and stored in a look-up-table for reference while the filter is operating.

Consider the ideal, or brick wall, digital low-pass filter with a cutoff frequency of ω_0 rad s^{-1}. This filter has magnitude 1 at all frequencies less than ω_0 and magnitude 0 at frequencies between ω_0 and the Nyquist frequency. The impulse response sequence $h(n)$ for a filter normalized for frequencies between 0 and π is

$$\begin{aligned} h(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} H(\omega)e^{jw\omega n}d\omega \\ &= \frac{1}{2\pi}\int_{-\omega_0}^{\omega_0} e^{j\omega n}d\omega \\ &= \frac{\omega_0}{\pi}\operatorname{sin}c\left(\frac{\omega_0}{\pi}n\right) \end{aligned} \tag{23-59}$$

This filter cannot be implemented as an FIR since its impulse response is infinite. To create a finite-duration impulse response, we truncate it by applying a window. By retaining the central section of impulse response in this truncation, you obtain a linear phase FIR filter. The length of the filter primarily controls the steepness of the cutoff, while the choice of window function allows you to trade off between pass band and stop band ripple, Fig. 23-23.

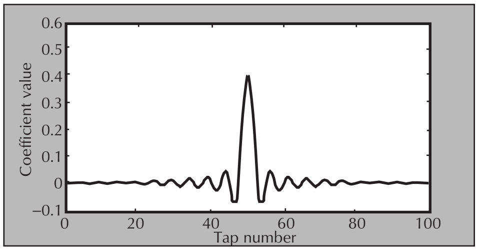


Figure 23-23. Coefficients of a 100-tap low pass filter at 0.2 times the sample rate.

23.5.1.2 FIR Length

The required number of taps (N) in an FIR filter at a sample rate (f_S) for a given transition band specified by a width (f_t) and attenuation in dB(A) can be estimated as

$$N = \frac{f_5 A}{f_t 22} \qquad (23\text{-}60)$$

As an example, we can calculate how may taps a 100 Hz, fourth order high-pass filter in a 48 kHz system would use. Fourth order gives a roll-off of 24 dB per octave, so the response will be 24 dB down by 50 Hz. The transition band is for a 20 dB minimum attenuation in the stop band and therefore is 50 × 20/24 or 42 Hz wide and the desired attenuation is 24 dB so the equation gives us 48,000 × 20 × /(42 × 22) = 1049 taps. This is a very long filter and introduces 520 samples or 10.8 ms of delay at the 48 kHz sample rate.

If we consider the same example, but for a 1000 Hz cutoff frequency, everything scales by a factor of 10, giving a much more acceptable filter length of 105 samples with a delay of 1.1 ms. This illustrates the limitation of using FIR filters for low frequencies, Fig. 23-24.

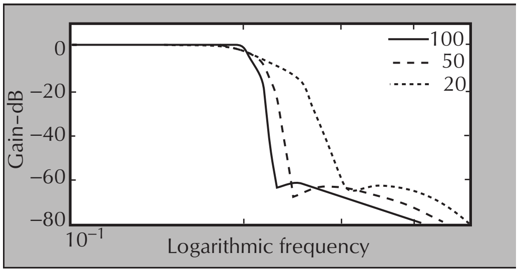


Figure 23-24. Increasing filter steepness with number of FIR taps.

23.5.2 IIR Filters

There are many possible configurations of *infinite impulse response* (IIR) filters, two of which are shown in Figs. 23-25 and Fig. 23-26. They show the direct form of a biquad filter in which the input and output samples are passed into the delay line. The transpose form has a sum between every delay and scaled copies of the input and output samples are inserted into the delay line. Direct form I is better suited to fixed point implementation where it is important that the delayed terms maintain as much precision as possible.

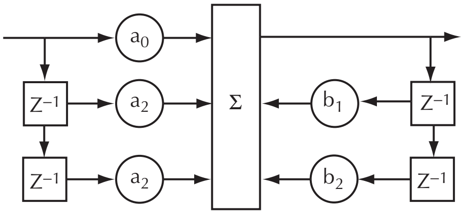


Figure 23-25. IIR Filter implementing a bi-quad section in Direct Form I.

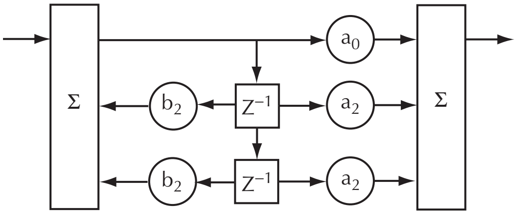


Figure 23-26. IIR filter implementing a bi-quad section in Direct Form II.

The biquad IIR filter is a second-order filter, and forms the most common basis for higher-order IIR filters. This form fits well with the transfer function equations such as the Butterworth polynomials in Table 23-1. The feedback coefficients correspond to the poles of the filter and the direct coefficients correspond to the zeros. Each section can be represented as a short FIR filter, but unlike the direct implementation of an FIR, these sections can have complex coefficients, even if the input and output are to be real only.

23.5.2.1 Calculation of Coefficients from Poles and Zeros

IIR filters are designed in terms of the Z transform. In this transform, the time domain representation of the filter is used and the notation z^{-1} is used in place of the common exponential terms in the discrete Fourier transform:

$$z^{-n} = e^{j\omega n\Delta t} \tag{23-61}$$

This gives us the expression for the *Z* transform:

$$H(z) = \sum_{n=-\infty}^{\infty} h[n]z^{-n} \tag{23-62}$$

The biquad has two poles in the denominator and two zeros in the numerator. It may be expressed in the factored form as:

$$H(z) = \frac{G(z - r_{01}e^{-jq_{01}})(z - r_{02}e^{-jq_{02}})}{(z - r_{p0}e^{-jq_{p0}})(z - r_{p1}e^{-jq_{p1}})} \tag{23-63}$$

where,
G is the gain,
r_0 denotes the real part of the zero location,
q_0 denotes the imaginary part of the zero location,
r_p denotes the real part of the pole location,
q_p denotes the imaginary part of the pole location.

Table 23-3 lists the equations for the individual coefficients for a purely real implementation of a biquad filter, given the locations of the poles and zeros.

Table 23-3. Relation of Biquad Coefficients to Pole and Zero Location

Zeros	Poles
$a_0 = 1$	
$a_1 = -2r_o cos(q_o)$	$b_1 = -2r_p cos(q_p)$
$a_2 = r_o^2$	$b_2 = r_p^2$

23.6 Equalizers

Equalizers are devices or components that are designed to compensate for undesirable characteristics in the magnitude or phase response of another part of the system and thus make the response equal again. Equalizers consist of filters implemented in such away as to provide control over the frequency response in terms of how the operator thinks of the response curve that they are trying to recreate. Equalizers give control over one or more of the parameters that affect the response over the audio range, usually 20 Hz to 20 kHz, and ideally do so such that the parameters do not interact. Controls are arranged in terms of center frequencies, bandwidths, and gains rather than actual circuit values that control these things. This means that often the controls are dual-ganged so that the ratio of two resistor values may be kept constant while their absolute values are changed.

23.6.1 Tone Control

The simplest form of equalizer is the *tone control* as used on portable radios. The control only acts to attenuate the high frequency. Another version of this type of equalizer that is becoming more prevalent than the tone control is the *bass boost,* which as its name suggests acts as the exact opposite to add a controlled gain to the low frequencies.

The tone control circuit shown in Fig. 23-27 includes transistor-based buffer amplifiers around the passive filter section in the middle. This allows the operation of the equalizer to be independent of source and load impedances.

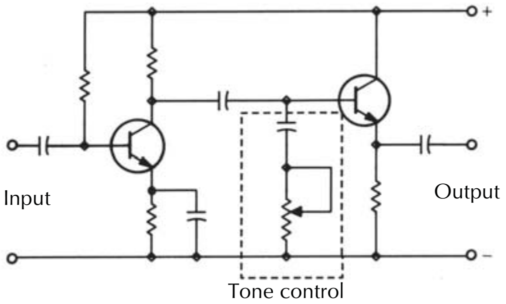


Figure 23-27. Simple low-pass tone control

23.6.2 Graphic Equalizers

A *graphic equalizer* is used to shape the overall spectrum of program material. The term *graphic* refers to the way that the controls are set out on the front panel such that the positions of the slider controls draw the desired frequency response. Graphic equalizers typically use ⅓-octave band filters but may be constructed with any spacing. The ⅓-octave refers to the spacing between adjacent filters and not necessarily the width of the filter.

A graphic equalizer is constructed using a series of filters with fixed frequency and width. The centers of the filters are typically on the ISO preferred frequencies rather than the mathematically correct ⅓-octave spacing. This means that in order to cover the spectrum completely, some of the filters must have different widths. The output of each filter is added to the original signal to a degree controlled by a slider control. The levels add together and are prone to producing a ripple in the response between the centers. In Fig. 23-28, the four sliders for 800 Hz, 1000 Hz, 1250 Hz, and 1600 Hz were set to +5 dB. The overall peak is greater than desired and a ripple of 2 dB is induced across the band.

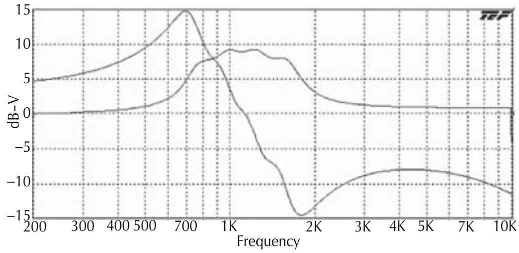


Figure 23-28. Magnitude and phase response of a graphic equalizer.

23.6.2.1 Transversal Equalizers

Fig. 23-28 shows an example of how graphic equalizers based on tuned filters exhibit ripple in the response when groups of adjacent controls are used. The actual response that we were trying to create would have been better achieved using a single filter as in Fig. 23-29. The Transversal *equalizer* configures as a graphic equalizer produces ripple-free response for any equal or flat setting of the controls. It produces minimum phase response curves and avoids phase mismatch anomalies at the band edges that can be a problem in other equalizers. The response curve is mathematically a best match for the desired response.

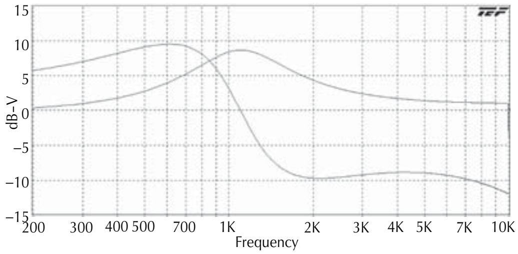


Figure 23-29. Magnitude and phase response of a single bi-quad filter.

The FIR filter discussed previously is a digital implementation of a transversal filter. Whereas a conventional tuned filter operates in the frequency domain, a transversal filter operates in the phase or time domain. If a unity gain all-pass circuit stage, Fig. 23-19, is substituted for each Z^{-1} delay element in Fig. 23-22, an analog transversal filter is created. The coefficients are implemented by summing the outputs of the successive delays via different weighting resistors to a summing amplifier.

23.6.3 Parametric Equalizers

Parametric equalizers allow adjustment of the filters in term of the three main parameters that define a filter.

- The boost or cut in dB.
- The center frequency.
- The bandwidth or *Q*.

It is difficult to make a parametric filter that provides completely independent control over all three parameters over a wide frequency range. Several filter components have to be varied with one control. For this reason, parametric equalizers sometimes have one of the controls as a multiposition switch instead of continuously variable. This allows a band of calibrated components to be switched into place rather than having to worry about how variable component values track.

Parametric equalizers are always active and typically there are several second-order sections in a unit. Each band's center frequency is adjustable over a limited frequency range so that the parameters' independence can be maintained. This means that each section in a unit typically covers a slightly different frequency range, each section having a ratio of between 10:1 and 25:1 between the highest and lowest center frequency. The lowest band will adjust down to 20 Hz and the highest band up to 20 kHz. Each section will typically provide more scope for cutting levels than for boosting. Typical boost level is up to 15 dB while the available cut may be down to −40 dB. The bandwidth or *Q* is not consistent in its labeling between manufacturers. Some specify bandwidth in Hz, some specify *Q*, and others specify octave fraction. In terms of *Q*, the range of this control will typically be between 0.3 and 3, with the critically damped value of 0.707 being in the center position of the control.

An overall gain is usually provided to help maintain the average level and to maximize headroom by avoiding clipping.

23.6.3.1 Semi-Parametric Equalizers

A reduced version of the parametric equalizer is commonly found on mixing consoles. This is the *semi-parametric* or *swept frequency equalizer*. This type has only the center frequency and cut or boost controls. The *Q* is usually set to be a midrange critically damped value but can also be configured so that the Q varies with gain.

23.6.3.2 Symmetric or Asymmetric Q

Straightforward designs produce *constant Q* filters that have the same Q for any amount of boost or cut. If the frequency response curves for the same amount of boost as cut are mirror images of each other across the unity gain axis, the response characteristic is called *reciprocal* or *symmetrical*. This means that the bandwidth of frequencies affected when boost is applied is greater than that affected when cut is applied. Fig. 23-30 shows that in the symmetrical response, the cutoff frequency in attenuation mode F_c is less than that in boost mode F_b.

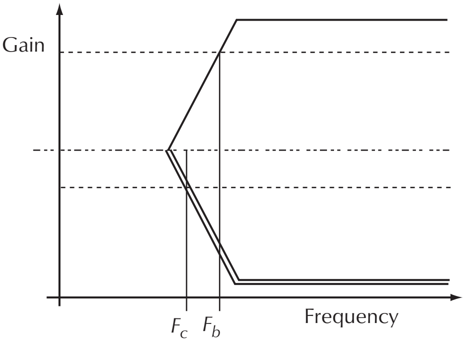


Figure 23-30. Symmetrical response with different bandwidth in cut and boost.

This is not always the most musically useful response. It is more common in spectrum shaping to want to gently apply boost to a broader region. Boosting a narrow region tends to lead to instability. At the same time, it is more useful to be able to notch out a fairly precise frequency, without removing a large portion of the surrounding spectrum. For this reason, equalizers tend be designed so that the bandwidth increases with gain.

23.6.4 Programmable Equalizers

All types of equalizers can be programmable. In digital equalizers, the filter coefficients are stored in memory and may be recalled or modified at will. Unless a digital equalizer implements only a fixed set of coefficients it is inherently programmable.

In programmable analog equalizers, a digital control system is used to physically manipulate the analog filters. This can be either by controlling switches that swap components in or out of the circuit, or by using voltage-controlled gain to alter the filter's response. In the case of switched capacitor filters, the digital control system can adjust the filters by manipulating the switching frequencies to adjust the equivalent resistor values and thus the filter characteristics.

23.6.5 Adaptive Equalizers

The *adaptive equalizers* have long been used in communications systems for multipath echo cancellation. They are the ultimate equalizers for sound systems that must adapt to acoustic conditions that may change at any time. A common example of an adaptive equalizer in sound reinforcement is a feedback suppressor. In this application, the equalizer monitors the signal passing through it for the characteristic exponential increase in level of a frequency that is associated with feedback buildup. When this increase is detected, a very narrow and deep notch filter is placed at that frequency to suppress the feedback. This can typically operate in a fraction of a second such that you were unaware that the event occurred.

References

1. J. Blauert, and P. Laws, "Group Delay Distortions in Electroacoustical Systems," *JASA*, Vol. 63, no. 5, pp. 1478–1483 (1978).
2. S. Linkwitz, and R. Riley, "Passive Crossover Networks for Non-coincident Devices," *JAES*, Vol. 26, no.3, pp. 149–150 (1978).

Chapter 24

Delay

by Steven McManus

24.1 Delay 805
  24.1.1 Comb Filter 805
  24.1.2 Directional Perception 805
24.2 Uses of Delay 806
  24.2.1 Delay in Loudspeaker Systems 806
  24.2.2 Setting Delay Times 807
  24.2.3 Reverberation Synthesis 808
  24.2.4 Delay-Based Effects 809
24.3 Implementations 809
  24.3.1 Small Delays 809
  24.3.2 Acoustic Delay Methods 809
  24.3.3 Tape Delay. 810
  24.3.4 Analog Shift Register Delays 810
  24.3.5 Digital Delays 810
24.4 Sampling in time 811
  24.4.1 Aliasing 811
  24.4.2 Capturing a Sample 812
  24.4.3 Errors in Sampling Magnitude 812
24.5 Analog-to-Digital Conversion 813
  24.5.1 Pulsed Code Modulation 813
  24.5.2 Delta Modulation. 813
  24.5.3 Sigma-Delta Modulation 814
  24.5.4 Decimation 814
References 815

This page intentionally left blank

24.1 Delay

Delay is relative. For a delay to have an effect on a sound it must be heard in conjunction with the original, nondelayed sound. There are two ways that this can occur. A single sound can arrive at the listener via two different length paths, such as a direct sound and a reflected sound, or two signals with different delays can be added electrically and then heard from a single location, Fig. 24-1.

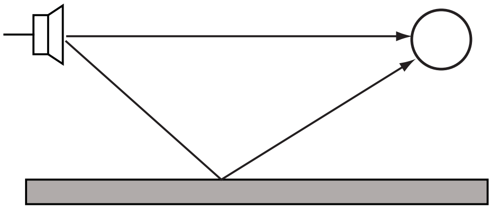


Figure 24-1. Different sound paths through air.

24.1.1 Comb Filter

Two copies of the same signal at different delay times combine to add or subtract depending on the relative phase of each frequency as shown in Fig. 24-2. If the waves are a whole period apart, they combine to give a peak in level, if they are half a period apart, they cancel out to an extent controlled by their relative levels. This effect sets up a comb filter, so named for its appearance on a frequency plot as shown in Figs. 24-3 and 24-4. The series of peaks in the response fall first at dc, then at every frequency whose period is equal to an integer multiple of the delay time. The cancellation notches occur at the exact midpoints between these frequencies.

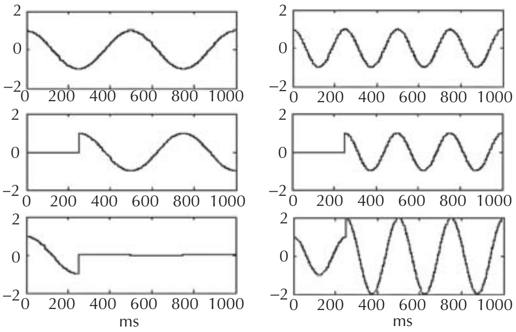


Figure 24-2. Effects of adding signals of different frequencies with the same delay.

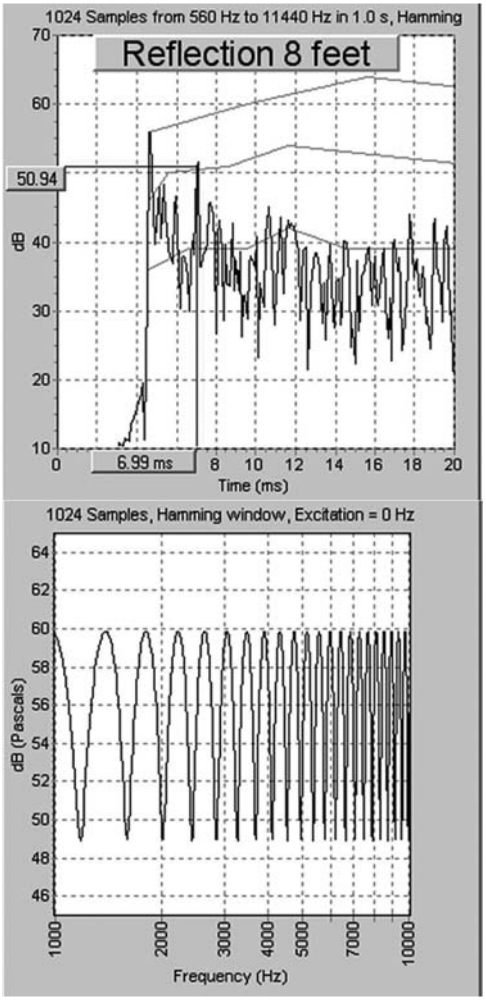


Figure 24-3. A reflection that will alter the perceived direction of the sound.

24.1.2 Directional Perception

In the case of sounds traveling through the air, the path lengths with their corresponding travel times are different for every point in space, resulting in a different comb filter for every location.

The brain uses the results of the different comb filters that are in effect at the location of each ear and combines this information in conjunction with the arrival times, relative levels, and directional filtering due to the shape of the pinnae to determine the originating direction of a sound. Other cues such as the ratio of direct to reverberant energy are used to help determine distance.

A completely dry sound heard in a set of headphones will appear to originate inside your own head. Gradu-

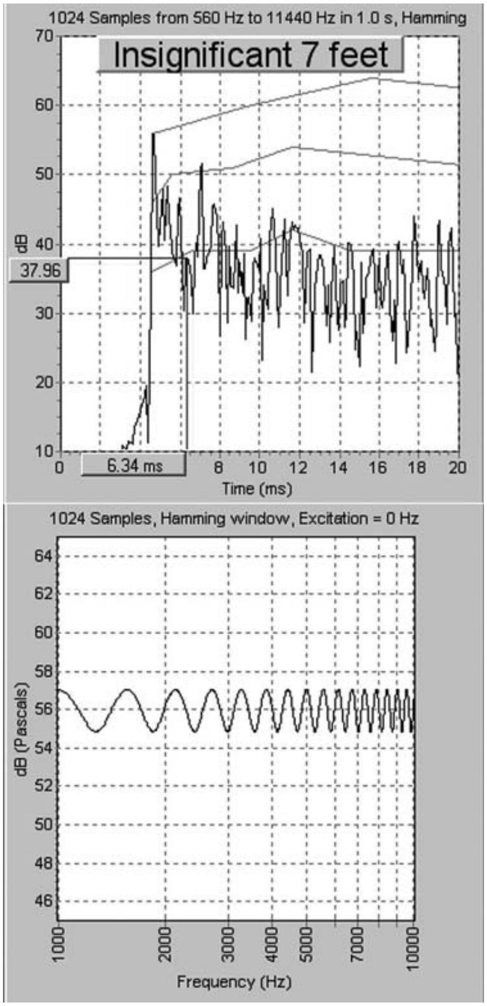


Figure 24-4. A reflection that is too low in level to affect perception of the sound.

ally adding reverberation to it will make the sound appear to move away out in front of you. The sound can be made to appear to move from side to side by alerting the relative levels in each ear, as is commonly done in a pan control, but the same effect can be achieved by altering the relative delay of the dry sound to each ear. The reasons that this delay technique is not commonly used are that the level control is much simpler to implement and the result is compatible with monaural reproduction when the left and right channels are summed.

A sound is perceived as originating in the location at which it was first heard. This is generally the correct location as the direct sound will always arrive before any reflected sounds. The same sound coming from a second location will be perceived in different ways depending on its timing and level relative to the first:

- If the second sound is more than 30 ms after the first it will be heard as a distinct echo.
- If the second sound is more than 10 dB louder than the first, it will be heard as a distinct echo.
- If the second sound is within 10 dB and less than 30 ms after the first, it will cause an image shift in where the source location is perceived.
- If the second sound is more than 10 dB below the first, it will contribute to the spatial feel of the sound but will not be heard as a distinct sound or alter the apparent location of the first.

These rules of thumb are approximations of the psychoacoustic effects in operation. The perception curves are more complex than the rules of thumb suggest. The actual values are plotted in Fig. 24-5 and tabulated in Table 24-1.

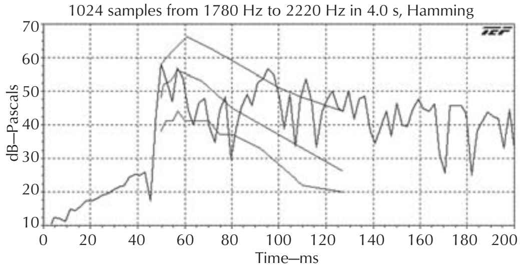


Figure 24-5. Perception curves.

24.2 Uses of Delay

Delay is sometimes useful. It should also be noted that there can be undesirable delays in a system. This is particularly true with digital recessing equipment where there is always a conversion delay in and out of the processor plus any processing delay. It is not uncommon for processors to have a minimum delay of a few milliseconds, and these delays should be considered when calculating the amount of delay that you actually want to use.

24.2.1 Delay in Loudspeaker Systems

A sound amplified through a loudspeaker system will be subject to image shifts and audible echoes only if there is a reference point against which to judge it. This is usually the case in sound reinforcement with the original sound being the reference or in multiple loudspeaker setups where another loudspeaker can act as the reference.

Table 24-1. Perception Curves of Figure 24-5 Tabulated

Echo	Image shift	Spacious	No effect
Ms after direct	dB	dB	dB
0	0	–10	–20
1		–6	
2			–17
4	4	–5	
5			–17
7		–2	–14
10			–17
11	8		
17	6	–5	
20	5		–17
25			–21
30		–13	–21
40			–25
50	–7		
60	–10		–36
77	–14	–32	–38

It is not generally desirable for loudspeakers in a system to appear to be generating echoes as this will have a detrimental effect on the intelligibility of the system. Whether the image shift effects are important depends on the application. In a stage system, it is desirable to have the apparent sound source at the stage, regardless of the placement of the loudspeakers. In a distributed announcement system, the creation of a coherent source image is not as important as the intelligibility.

Sound travels at 334 m/s or 1130 ft/s. A sound traveling 33 ft will be delayed by 30 ms, so with sound sources greater than 33 ft apart, delay should be used to avoid the creation of echoes.

24.2.2 Setting Delay Times

In Fig. 24-6 the sound from the source, a person talking, is to be augmented by a loudspeaker and the apparent source of the sound is to be kept on the stage. To achieve this, the sound from the source must arrive at the listener before the sound from the loudspeaker. The time taken for the signal to arrive at the listener from the loudspeaker is a combination of the distance traveled in air from the loudspeaker to the listener and the negligible time taken for the signal to arrive electrically at the speaker. We must delay the signal to the loudspeaker by an amount that allows the direct sound traveling more slowly though the air to catch up and overtake the sound from the loudspeaker. The delay should slightly exceed the time taken for the sound to travel the difference in distance between the source and the loudspeaker so that the direct sound will be heard first and localized to the source. The loudspeaker can then add up to 10 dB of level 5 to 10 ms later to increase the level of the sound without changing its apparent position.

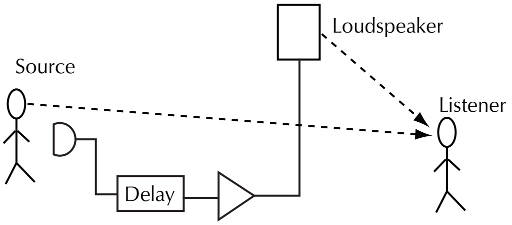


Figure 24-6. A delay in the sound system corrects for the differences in path length between the source and listener and the loudspeaker and listener.

A graphical method for setting delays is shown in Fig. 24-7. The positions of the source and loudspeakers are plotted and a series of concentric circles drawn around them at 30 ms (33 ft) intervals. The SPL level from the polar response pattern of the loudspeaker can also be plotted, but for simplicity in this example, omnidirectional sources are used where the level decreases by 6 dB per doubling of distance.

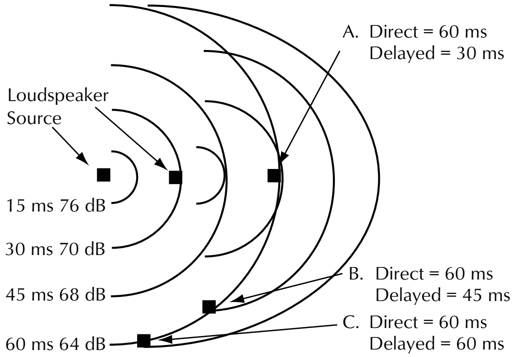


Figure 24-7. Graphical method of setting delays.

If we look at point A, where source and loudspeaker are in a direct line, the time difference is 30 ms. If we add a small amount to this to allow the direct sound to be heard first, we come up with a delay setting of

35 ms. We can now analyze the sound at each of the three points:

1. The loudspeaker is 6 dB louder than the source and 5 ms later. The overall sound is heard as originating at the source but 6 dB louder.
2. The loudspeaker is 2 dB louder than the source and 20 ms later. The overall sound is heard as originating at the source but 2 dB louder.
3. The loudspeaker is the same level as the source but 35 ms later. At this point the delay is too long and a distinct echo is heard.

In reality, the coverage pattern of the loudspeaker should be chosen to ensure that the level of the sound is sufficiently attenuated outside the area where the delay works effectively.

24.2.3 Reverberation Synthesis

Reverberation is the result of many reflections of the original sound. The general pattern of events, as shown in Fig. 24-8 is that there is first a direct sound, followed by a short gap, referred to as the initial time gap (ITG). Next come the first distinct early reflection echoes caused by sound bouncing off surfaces near either the source or the listener. Thereafter the reflected sounds start to generate their own second, third and higher-order reflections and the energy level settles down to a constant decay rate. This decay rate is related to the distances traveled and the amount of absorption in the room.

Delay is used as the basis for reverberation synthesis because it provides a convenient method for storing the signal and releasing it at a later time, much as reflections from the surfaces of a room arrive at the listener at a later time than the direct sound. Typical applications for synthetic reverberation include the enhancement of program material in the production of recordings, the introduction of special effects in live entertainment productions, and compensation for poor or lacking natural reverberation in entertainment spaces.

Requirements for good reverberation synthesis are essentially the same as for an acoustically well-designed hall. There are many parameters that need to be considered to help achieve realism in reverberation simulation.

- **Distance from Source.** The perception of distance is controlled primarily by the relative energy levels of the direct components and the decay components.
- **Room size.** The perceived room size is controlled by the delay time from the first grouping of direct sound and early reflections to the start of the decay tail and by the length of the decay tail. The requirements for the decay tail are the same as for an acoustically well- designed room. A relatively smooth decay rate is desirable, with longer decay times at lower frequencies than at higher frequencies to simulate the high-frequency losses as sound travels through the air.
- **Brightness.** The spectral balance of the decay tail determines the character of the reverberation. A lot of high- frequency roll-off simulates a room with a lot of absorption from carpets, curtains, or designed absorption devices and gives a dark sound. Less height frequency roll-off simulates a hard-surfaced room such as the inside of a stone church, giving a bright sound.
- **Character.** The smoothness of the decay determines the character of the sound. A room with large opposing flat surfaces will exhibit a *flutter echo* where the sound bounces back and forth between the walls with little diffusion. A room with more architectural features or multiple surfaces will tend to scatter the sound more, creating a denser and more evenly distributed decay.

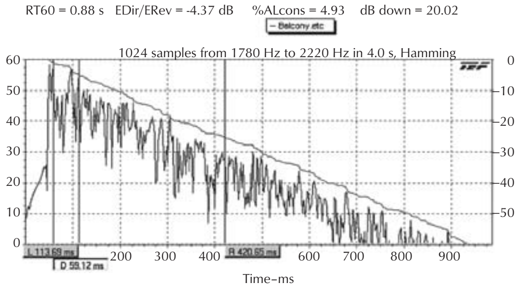


Figure 24-8. Energy Time Curve showing the delay of sound in a room

- **Envelopment.** The sense of the reverberation coming from all around you rather than a specific location is controlled, making different patterns for the different playback channels. This can be effectively achieved using two-channel stereo as well as in systems with multiple dedicated speakers. The most important differences are in the pattern of the early reflections. The decay tail portions should keep the same with the direct to reverberant energy ratio and decay time but can be randomized to produce a denser sound field. Portions of the randomized signal can be altered in their frequency response to mimic the ear's nonuniform response to sounds from behind, further increasing the sense of envelopment.

A tapped delay a shown in Fig. 24-9 is suitable for creating early reflections. Delays $T_1 \ldots T_N$ are unequal in length and in the range of 10 ms to 30 ms with amplitudes set by $g_1 \ldots g_N$ as appropriate for the character of the room. In Fig. 24-10, a reticulating path is provided via g_f that produces the exponentially decaying portion. This could simply be fed into the start of the reflection generator, but a more satisfactory result is obtained by using a separate decay section, where the delay taps are set more densely and may be over a wider range, typically between 5 ms and 100 ms. The delay tap times should be chosen to not be harmonic products of each other to minimize the buildup of standing waves and comb filters. Any gain product in the reticulating path must be less than one, otherwise, the sound will exponentially increase until distortion occurs.

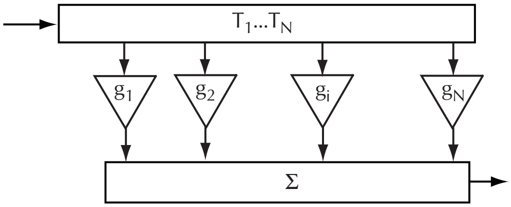


Figure 24-9. A delay with multiple taps for creating early reflections.

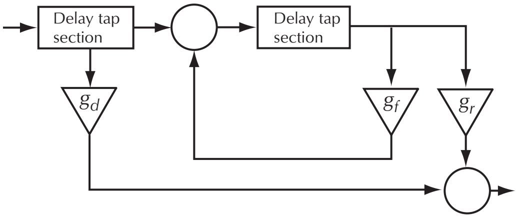


Figure 24-10. Delay section with feedback for producing a decay tail

24.2.4 Delay-Based Effects

Flange is an audio effect caused by mixing an original (dry) copy of a sound with a delayed (wet) copy. The amount of delay is varied over time, creating a varying pattern of comb filters that sweep up and down through the audio spectrum. Chorus is used to make one voice or instrument sound like many and has the same topography as a flanger, but with longer delays.

24.3 Implementations

The implementation of a delay requires some means of storing the signal and then releasing it after a controlled period of time. This can be done either by storing a continuous record of the sound or by breaking it up into samples that are stored separately. Some preparation of the signal is usually required to make it compatible with the chosen storage medium and method and may involve some postprocessing to restore the stored signal to a usable form.

24.3.1 Small Delays

Small delays may be realized using the phase shift characteristics of an all-pass filter. Such a circuit, illustrated in Fig. 24-11, is limited to delays of the order of less than a wavelength of the highest frequency. These sections may be chained together to produce longer delays but become impractical for delays longer than a few milliseconds. This method is sometimes used in active crossover systems for a loudspeaker. The delays needed to time-align drivers within a cabinet are small and fixed, and the circuit may be easily combined with the frequency filtering requirements.

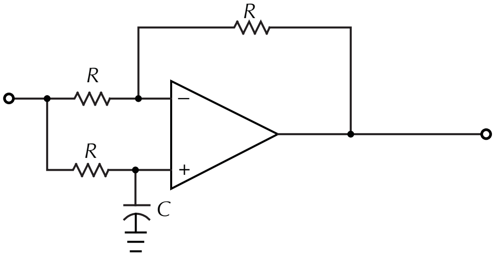


Figure 24-11. An all-pass amplifier having phase shift proportional to frequency and exhibiting a small amount of delay.

24.3.2 Acoustic Delay Methods

One way to implement a long delay is to use the speed of sound and send the signal to be delayed through a fixed air path, such as a tube with a loudspeaker at one end and a microphone at the other. For such a device to work effectively, the tube must be damped to prevent internal reflections and have an absorber at one end to prevent the establishment of standing waves. This type of system has many disadvantages as the system becomes very large for any useful delay time. The frequency response changes with tube length due to the damping material and the signal attenuates as it travels, meaning that large amounts of gain are required. The large gain in turn leads to the requirement that the tube must be mechanically isolated from vibration and outside sounds to prevent these from being added to the delayed sound.

24.3.3 Tape Delay

A more practical early implementation method for continuous delay was to use a magnetic tape loop. An example of such a device is shown in Fig. 24-12. The sound is recorded onto the tape at the record head and then is read back by one or more playback heads. The tape then passes an erase head and loops back to the start. The delay time is given as

$$Time = \frac{distance\ between\ heads}{tape\ speed} \tag{24-1}$$

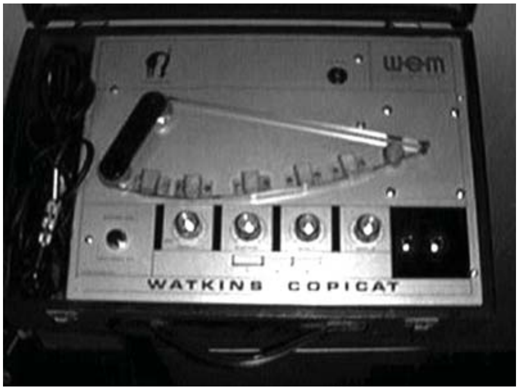


Figure 24-12. Tape loop delay system.

Only the length of the tape limits the maximum delay time. The performance of this system depends on the quality of the recording system. Dynamic range, frequency response, and SNR are affected by the tape speed and track width. These parameters may be improved by using the usual tape recording tricks such as compression and various forms of preemphasis/deemphasis noise reduction. Such systems need regular maintenance including head cleaning and replacement of the tape to maintain optimum performance.

24.3.4 Analog Shift Register Delays

Analog shift registers as illustrated in Fig. 24-13 appear in two forms, the bucket brigade and the analog charge coupled device (CCD). They both operate in a very similar manner and differ at the silicon level in the type of switches: a metal oxide-semiconductor capacitor (MOSC) structure for a CCD and a metaloxide–semiconductor junction FET capacitor (MOSJC) structure for the bucket brigade.[2] They are classed as shift registers because they move single samples of a signal in the form of an electrical charge from one stage to the next in response to timing signals. The delay (T) of a shift register is proportional to the number of register elements (N) and inversely proportional to the frequency (f_s) of the timing signal.

$$T = \frac{N}{f_s} \tag{24-2}$$

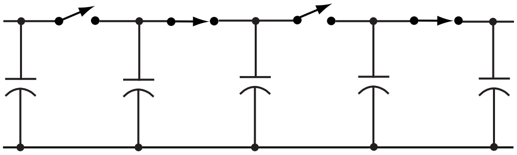


Figure 24-13. Bucket brigade: switched alternately open and close to hand charges along the line.

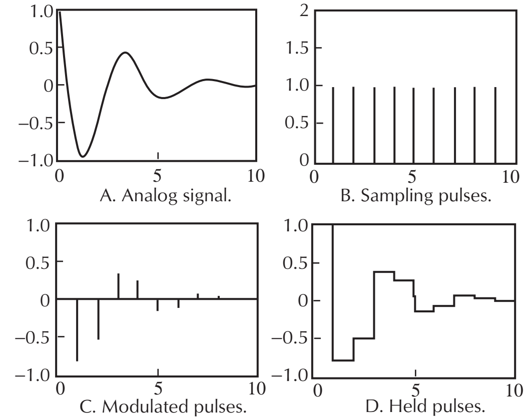
The term *charge transfer device* (CTD) has been applied to both the bucket brigade and CCD-based structures of an analog delay. The term *CCD* has become colloquially associated with a type of light-sensitive array used in cameras but actually refers to the method used to read the information off these devices. The idea of a CTD is that it stores a sample of analog information as a packet of charge on a capacitor and, under control of a timing signal, transfers it to the next storage site. All the requirements of sampling theory for audio-frequency band limiting should be met. The performance parameters for a CTD include transfer efficiency (ε), the fraction of charge left behind in each transfer; the leakage of charge from a cell during the holding period; and the leakage of charge into a cell due to semiconductor thermal effects. Taken together, these effects degrade the SNR of the signal as it passes through the CTD and also lead to distortion due to the nonlinear nature of the leakages. The practical use of CTD is limited to applications requiring less than 100 ms of delay or longer where SNR and distortion may be tolerated. CTDs have largely fallen out of use as delay lines because digital systems have become cheaper.

24.3.5 Digital Delays

Digital delay operating principles have undergone a number of important changes since they were first used in sound systems. Probably the most significant is in the type of storage or memory. Shift registers were almost universally used in the first commercially produced units. Digital shift registers are conceptually similar to analog CTD devices with the important advantage that only the presence or absence of a charge carries the significant signal information. Now random access memory (RAM) provides flexibility and economic tradeoffs for design. Until recently, the cost of memory was the dominant factor in delay design considerations. Currently, with the trend for DSPs to include large amounts of on-board memory, the systems have vastly reduced in cost and now the dominant cost factor is in the A/D and D/A converters.

24.4 Sampling in time

Both analog CTD delays and digital delays rely on breaking the delayed signal up into discrete samples. These samples are created by looking at a signal's amplitude at regular intervals and disregarding its amplitude at all other times. The procedure is shown in Fig. 24-14. The sequence of pulses (B) controls a switch that turns on the signal (A) for a brief instant, then disconnects it for the remainder of the sampling period. The result is an amplitude-modulated pulse train (C) where each pulse has amplitude equal to the instantaneous signal value. According to the sampling theorem, a continuous bandwidth-limited signal that contains no frequency components higher than a frequency f_c can be recreated if it is sampled at a rate greater than $2f_c$ samples per second. This rate is called the *Nyquist frequency*. Since the real world never completely satisfies theoretical conditions, sampling frequencies are usually chosen to be higher than $2f_c$. Thus, 20 kHz bandwidth delays will typically be found with a sampling frequency of 48 kHz rather than the bare minimum of 40 kHz.



A. Analog signal. B. Sampling pulses.

C. Modulated pulses. D. Held pulses.

Figure 24-14. The process of sampling a signal.

24.4.1 Aliasing

Sampling of the audio signal is a form of modulation. Modulation of a bandwidth-limited signal with an upper frequency of f_c, by the sampling frequency f_s produces additional copies of the original spectrum centered on frequencies f_s, $2f_s$, $3f_s$, etc. If the sampling frequency is not high enough or the bandwidth is not adequately limited, part of the spectrum centered on f_s will fold over into the original signal spectrum as in Fig. 24-15. The fold-over components become part of the signal in the recovery process, producing unwanted frequencies that cannot be filtered out.

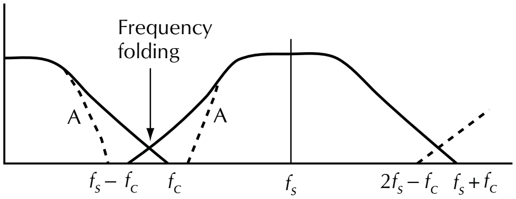


Figure 24-15. Frequency spectrum folding over around the sampling frequency.

An example of the effect of aliasing can be seen on moving wagon wheels in a movie that appear to reverse direction. The sampling rate of the film is lower than the rate at which individual spokes pass the top of the wheel. When the image is reconstructed, the spoke frequency has folded over and the wheel appears to move at a different rate. This phenomenon is known as *aliasing*. Fig. 24-16 shows that aliasing where the sample points lie have the same amplitude on two waveforms of different frequencies.

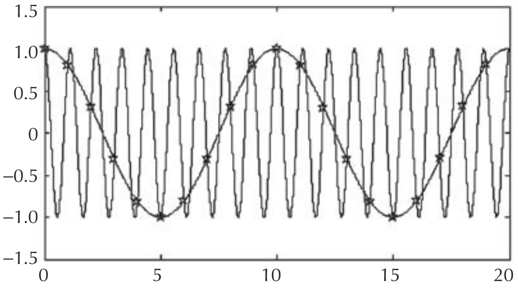


Figure 24-16. Aliasing.

Aliasing must be eliminated or at least largely reduced by selection of a high sampling rate and an adequately sharp *antialiasing* filter. At the output side, a similar low-pass *antiimaging* filter must be used to reduce the number of high-frequency glitches due to the switching at the sample rate.

The economics of delay design dictate that a relatively low sampling rate be used as this reduces the amount of storage that is required for a given length of signal. The number of storage locations required is the product of the sample rate and the length of the delay.

The required cutoff rate of the antialias filter is governed by the separation of the upper frequency f_c and the Nyquist frequency $f_s/2$. As these two frequencies become closer, the number of poles required in the filter increases, adding cost to the filter.

Antialias filters may be implemented as either analog or digital circuits. A digital antialiasing filter still requires some form of an analog antialias filter but relies on a high rate of *oversampling* to ease the design requirements. The digital filter does not have the demanding memory requirements of the delay line, so it can operate at a much higher sample rate than the storage section.

24.4.2 Capturing a Sample

A *sample and hold* circuit takes a very fast snapshot sample of the instantaneous voltage of an analog signal and then changes into a hold mode to preserve that voltage. A hold circuit forces the amplitude of the sample to have constant value throughout a sample period. In Fig. 24-14D, the sample amplitude is shown being set at the beginning of the sample period.

A basic sample and hold circuit is shown in Fig. 24-17. The signal amplitude is frozen for a brief period of time on a capacitor until the next sample period is initiated, at which time the new signal amplitude is transferred to the capacitor. The switch is momentarily closed, under the control of the sample pulse, and then reopened. The amplifier A_1 must have low-output impedance to make it capable of driving enough current to charge the capacitor to the appropriate voltage during the brief ontime of the sampling pulse. The output amplifier A_2 must have high-input impedance so as not to draw excessive charge from the capacitor as any leakage of current will cause a change in the voltage. The capacitor should also be lowleakage to help hold the voltage stable. An analog delay may be constructed entirely out of sample and hold circuits that transfer charge from one to another.

24.4.3 Errors in Sampling Magnitude

Any sampling system, digital or analog, will take a finite time to convert the input voltage into a form suitable for storage. This time is called the *aperture time* and relates to the amplitude resolution of the conversion. The sampling error ΔV is equal to the amount that the input voltage, V, changes during the aperture time t_α:

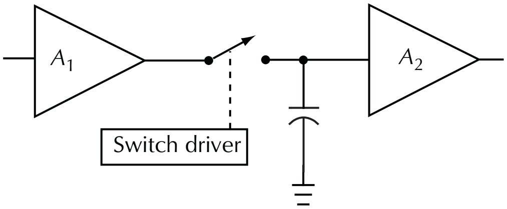


Figure 24-17. A commonly used sample-and-hold circuit.

$$\Delta V = t_\alpha \frac{dV}{dt} \tag{24-3}$$

For a sinusoidal input with peak amplitude A

$$\Delta V = t_\alpha \frac{d}{dt} A \sin \omega t \tag{24-4}$$

$$\Delta V = t_\alpha A \omega \cos \omega t$$

where,
ω is 2πf.

The rate of change of voltage is greatest at the zero crossing when $t = 0$

$$\Delta V = t_\alpha A \omega \tag{24-5}$$

Expressing this error, e, a fraction of full scale,

$$\begin{aligned} e &= \frac{\Delta V}{2A} \\ &= \pi f t_\alpha \end{aligned} \tag{24-6}$$

where,
V is voltage,
A is peak amplitude,
f is frequency,
t_α is aperture time.

As an example, a 20 kHz signal samples to a resolution of 16 bits (1 part in 65,536 or 0.0000152) requires an aperture time of 0.0000152/20,000π, or 0.24 ns. This is a very short time interval for an analog-to-digital converter to operate in. A sample and hold circuit is used to preserve the voltage long enough for the conversion to take place. The aperture time of the system becomes the open switch time of the sample and hold rather than the conversion time of the analog-to-digital converter (ADC).

24.5 Analog-to-Digital Conversion

Details for the large number of analog-to-digital conversion methods are outside the scope of this chapter, but the efficiency with which it is accomplished is so important to the success and acceptability of a digital delay or reverberation system that an overview of the common conversion principles is useful.

24.5.1 Pulsed Code Modulation

Pulsed code modulation (PCM) uses a number to represent the value of each sample. The continuously varying analog signal is divided up in time by *sampling* and divided up in amplitude by *quantization*.

The quantization resolution is defined by the number of bits used in the binary number and defines the amplitude resolution of the signal. The number of possible states for a number with n bits is 2^n. For a 16-bit number, there are 2^{16} or 65,536 different voltages that may be represented. For a 1 V peak-to-peak signal, this is equivalent to a 30 μV resolution. An error in the representation of the analog value results because there is a range of voltages that yield the same output code. This error is called the *quantization noise* and is given by

$$Q = \frac{A}{2^n} \tag{24-7}$$

where,

Q is the smallest analog difference that can be resolved by the converter,

A is the maximum amplitude,

n is the number of bits.

Another way of expressing this error is as the dynamic range of the converter.

$$\begin{aligned} DR &= 20\log 2^n \\ &= 20n\log 2 \\ &= 6.02n \end{aligned} \tag{24-8}$$

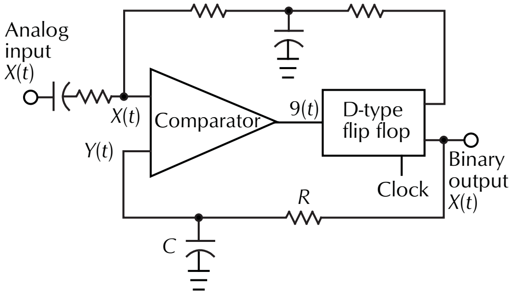
A 16-bit coding system will therefore have a dynamic range of 6.02^{16} = 96 dB.

The multibit binary word represents the amplitude of samples at regular intervals, usually in *twos complement* form. In this scheme the codes vary between 2^{n-1} and $2^{n-1} - 1$. The *most significant bit* (MSB) indicated the sign, with all negative values having MSB = 1. The code is often used in it fractional form, where the numbers represent values between –1 and 0.999.

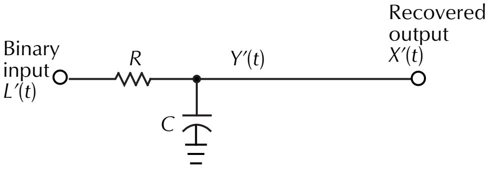
24.5.2 Delta Modulation

Delta modulation is based on whether the newest sample in a sequence is less than or greater than the last. A delta modulator produces a stream of single bits representing the error between the actual input signal and that reconstructed by the demodulator.

A simple delta modulator, as shown in Fig. 24-18, consists of three parts: a comparator whose output is high or low depending on the relative levels of the input signal (SST) and the reconstructed signal $y(t)$, a D-type flip-flop that stores the comparator output under control of a sampling clock, and a reference decoder that integrates the binary output to reconstruct the signal $y(t)$. The demodulator is a simple integrator circuit with the same characteristics as those used in the reference path of the modulator. It will reconstruct the reference signal $y'(t)$, which is a close approximation of the original input.



A. Coder.



B. Decoder.

Figure 24-18. A delta-modulator system.

The simplicity of the coding and decoding schemes has resulted in use of the delta modulator for communication and motor control applications. The simplest integrating network consists of a resistor and a capacitor, but the quantization noise from this is quite high so more practical systems use double integration in the reference path.

Distortion in delta modulation occurs if the rate of change in the input signal is greater than the maximum rate of change of the output of the integrator. The maximum rate at which a sinusoidal signal, $A\sin\omega t$, varies is

$A\omega$. The maximum required charge in voltage per sample is therefore

$$\Delta V = A\omega\Delta t \tag{24-9}$$

The value of ΔV relative to the maximum amplitude A defines the amplitude resolution of the system. The SNR is proportional to the sampling frequency and inversely proportional to the signal bandwidth. When presented with a fixed input, the delta modulator will hunt for the value by changing the output between 1 and 0 every sample. The resulting output is a tone of magnitude ΔV at the sampling frequency and is called the *idling noise*.

Delta modulation is more immune to errors in storage or transmission than PCM. A single-bit error in the output has a resulting error in the analog signal of ΔV. In a PCM system a single-bit error could cause an error of up to half the full-scale value. When compared to a PCM system in term of bits per second, delta modulation will have comparable dynamic range but a smaller frequency range. At lower bit rates, delta modulation can have a better SNR and dynamic range than a PCM system and this has implications for delay lines, where the total number of bits that must be stored can be reduced for the same quality of signal.

24.5.3 Sigma-Delta Modulation

By reorganizing the sequence of operations, the delta Modulator becomes a *sigma-delta modulator*. A first-order SDM as shown in Fig. 24-19, has one low-pass filer integrator in the signal path and a direct feedback path to a summing point that produces an analog error signal. The comparator of the delta modulator is replaced with a quantizer, which is a comparator against a fixed zero reference. Demodulation is accomplished using a low-pass filter as in delta modulation.

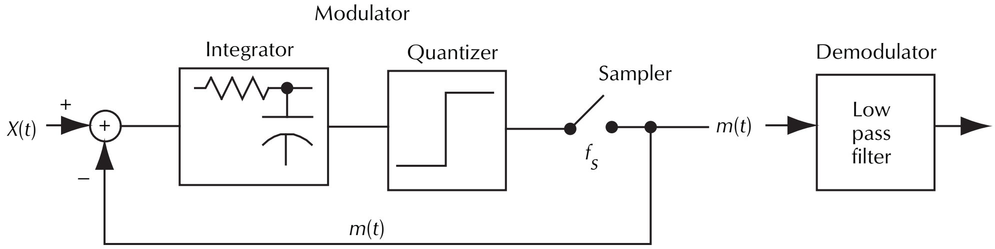
Both delta modulators and sigma-delta modulators use sampling frequencies much larger than the Nyquist frequency, typically of the order of 100 times. This places the quantization noise energy at very high frequencies where it can easily be removed by filtering.

24.5.4 Decimation

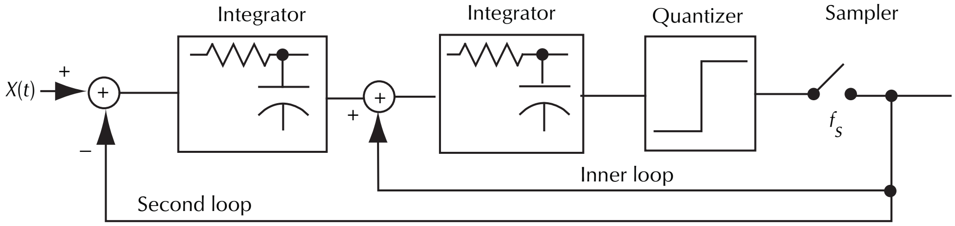
The process of sampling well above the required Nyquist frequency is called *oversampling*. The objective is to cause the modulation noise inherent in the sampling process to appear at frequencies further removed from the audio signal so that it can be more easily removed by filtering. The high SNR of an oversampled system can be preserved while reducing the overall bit rate by the process of *decimation*.

The quantization noise from PCM encoding decreases by 6 dB for every bit added but decreases by only 3 dB for every doubling of the sample rate. Decimation of the oversampled PCM data can result in a large reduction in the overall bit rate. Decimation filters are used to achieve this and can be thought of as performing an interpolation on the existing data to fill in the additional bits in the output.

A sigma-delta modulator can have as much as a 15 dB SNR improvement for a doubling of the sample rate. Decimation of the SDM signal can be used to convert the single bit data stream to a multibit PCM format suitable for storage in RAM and processing by DSP.


Modulator
Integrator
Quantizer
Sampler
Demodulator
X(t)
+
−
f_s
m(t)
Low pass filter
m(t)

A. First order.


Integrator
Integrator
Quantizer
Sampler
X(t)
+
−
+
f_s
Inner loop
Second loop

B. Second order.

Figure 24-19. Sigma-delta modulators.

References

1. H. Haas, "The Influence of a Single Echo on the Audibility of Speech," *J. Audio Eng. Soc.*, Vol. 20 (March 1972), pp. 145–159.

2. L. G. Heller, W. H. Chang, A. W. Lo, "A Model of Charge Transfer in Bucket Brigade and Charge-Coupled Devices," *IBM Journal of Research and Development*, Vol. 16, no. 2, p. 184, 1972.

This page intentionally left blank

Chapter 25

Consoles and Computers

by Steve Dove

25.1 Introduction . . . 825
  25.1.1 Console Development . . . 825
25.2 Auxiliary Paths . . . 826
  25.2.1 Monitoring . . . 826
  25.2.2 Prefade Listen and Audition . . . 827
  25.2.3 Overdubbing and Foldback (Cue) . . . 828
  25.2.4 Echo, Reverberation, and Effects Send . . . 828
  25.2.5 Communications (Talkback) . . . 829
  25.2.6 Combined Auxiliaries . . . 829
25.3 Stereo Consoles . . . 829
  25.3.1 Panning . . . 829
  25.3.2 Auxiliaries for Stereo . . . 829
  25.3.3 Multiple Effect Feeds . . . 830
25.4 Dawning of Multitrack . . . 830
25.5 Grouping and the Monitoring Section . . . 831
  25.5.1 Subgrouping and Output Matrices . . . 833
25.6 Console Design Developments . . . 833
  25.6.1 Transistors . . . 834
  25.6.2 Level Specifications . . . 834
25.7 Operational Amplifiers in Consoles . . . 835
  25.7.1 Slew-Rate Limitations . . . 836
  25.7.2 Bipolar Field Effect Transistors (BiFETs) . . . 837
  25.7.3 Discrete Operational Amplifiers . . . 838
  25.7.4 Instability . . . 838
  25.7.5 Phase Margin . . . 838
  25.7.6 Time-Domain Effects . . . 839
  25.7.7 Transient Intermodulation Distortion (TID) . . . 840
  25.7.8 Output Impedance . . . 840
  25.7.9 Compensating Op-Amps . . . 841
  25.7.10 Input Saturation . . . 842
  25.7.11 Front-End Instability . . . 842
  25.7.12 Band Limiting . . . 843
  25.7.13 Slew-Rate Effects . . . 844
  25.7.14 Device Idiosyncrasies and the Future . . . 844
25.8 Grounding . . . 844
  25.8.1 Wire . . . 845
  25.8.2 Earth Ground . . . 846
  25.8.3 Why Ground Anything to Earth? . . . 846
  25.8.4 Console Internal Grounding . . . 847

    - 25.8.4.1 Ground Follows Signal 847
    - 25.8.4.2 Ground Current Summing 848
    - 25.8.4.3 Typical Grounding Problems 848
    - 25.8.4.4 Ground Noise in Virtual-Earth Mixers 849
    - 25.8.4.5 Reactive Ground Effects 850
- 25.9 Signal Switching and Routing 850
  - 25.9.1 Relays 850
  - 25.9.2 Electronic Switching 851
  - 25.9.3 MOSFETs and CMOS 851
  - 25.9.4 Potentiometric Switching 853
  - 25.9.5 Minimizing Noise 853
  - 25.9.6 Practical Matrix Example 854
  - 25.9.7 Matrix Crosstalk 854
  - 25.9.8 16-Track or 32-Track Routing 854
  - 25.9.9 Processor Control 855
  - 25.9.10 Audio Path 855
- 25.10 Input Amplifier Design 858
  - 25.10.1 Power Transfer and Termination Impedance 859
  - 25.10.2 Optimizing Noise Performance. 860
  - 25.10.3 Noise Figure 860
  - 25.10.4 Inverse-Frequency Noise 860
  - 25.10.5 Optimum Source Impedance (OSI) 861
  - 25.10.6 Microphone Transformers 861
    - 25.10.6.1 Transformer Characteristics 862
    - 25.10.6.2 Common-Mode Rejection 863
    - 25.10.6.3 2Microphone Transformer Model 863
    - 25.10.6.4 Bandwidth 864
    - 25.10.6.5 Common-Mode Transfer 864
    - 25.10.6.6 Input Impedance 865
    - 25.10.6.7 Attenuator Pads 865
    - 25.10.6.8 Transformerless Front Ends 866
    - 25.10.6.9 Line Level Inputs 867
    - 25.10.6.10 Electronic Balanced Outputs 869
  - 25.10.7 A Practical Microphone-Amplifier Design. 871
    - 25.10.7.1 System-Level Architecture 871
    - 25.10.7.2 Input Coupling 872
    - 25.10.7.3 Line Level Input Facility 873
    - 25.10.7.4 Common-Mode Rejection Ratio 873
    - 25.10.7.5 Minimum-Gain Considerations 874
    - 25.10.7.6 The Limiter 874
    - 25.10.7.7 High-Pass Filters 875
- 25.11 Equalizers and Equalization 875
  - 25.11.1 Single-Order Networks 876
  - 25.11.2 Single-Order Active Filters. 876
  - 25.11.3 Changing Filter Frequency 878
  - 25.11.4 25.11.4 Reactance and Phase Shifts. 878
  - 25.11.5 Resonance 879
  - 25.11.6 Resonant Q 879
  - 25.11.7 Bandwidth and Q 879
  - 25.11.8 Creating Inductance 879

- 25.11.9 Gyrator Types 880
- 25.11.10 The Totem Pole 880
- 25.11.11 The Bootstrap 881
- 25.11.12 Constant-Amplitude Phase-Shift Network 883
- 25.11.13 Simulated Resonance 883
- 25.11.14 Consistent Q and Constant Bandwidth 883
- 25.11.15 Q, EQ, and Music 883
- 25.11.16 Bandpass Filter Development 884
- 25.11.17 Listening to Q 884
- 25.11.18 Push or Retard? 885
- 25.11.19 The Two-Integrator Loop 885
- 25.11.20 Stability and Noise Characteristics 887
- 25.11.21 Q and Filter Gain 888
- 25.11.22 High-Pass Filters 888
- 25.11.23 Second or Third or More Order? 889
- 25.11.24 Equalization Control 890
- 25.11.25 The Baxandall 890
- 25.11.26 Swinging Output Control 892
- 25.11.27 Swinging Input Control 892
- 25.11.28 A Practical EQ 894
- 25.11.29 The First EQ Stage 894
- 25.11.30 Second EQ/Line Amp 894
- 25.11.31 Low-Frequency Control 894
- 25.11.32 High-Frequency Control 894

25.12 Dynamics 895
- 25.12.1 Limiting 895
  - 25.12.1.1 Side-chain Time Constants 900
  - 25.12.1.2 Attack Time 900
  - 25.12.1.3 Release Time 901
  - 25.12.1.4 Compound Release Time Constants 902
  - 25.12.1.5 Multiband Processing 902
  - 25.12.1.6 Active Release Time Constants 903
  - 25.12.1.7 Hold or Hang Time 903
  - 25.12.1.8 Limiting and Compression for Effect 904
- 25.12.2 Gating 905
  - 25.12.2.1 Gating Feed-Forward Side Chain 905
  - 25.12.2.2 Subtractive Feedback Gate or Expandor 906
  - 25.12.2.3 Keying 906
- 25.12.3 Compression 906
- 25.12.4 Expansion 907
- 25.12.5 Feed-Forward VCA-Style Dynamics 907
  - 25.12.5.1 rms Detection 908
  - 25.12.5.2 Thresholding 908
  - 25.12.5.3 Ratio 908
  - 25.12.5.4 Side-Chain Sources 909
  - 25.12.5.5 Makeup Gain 909
- 25.12.6 A Practical Feed-Forward Design 909

25.13 Mixing 910
- 25.13.1 Virtual-Earth Mixers 910
- 25.13.2 Noise Sources 910

25.13.3 Radio-Frequency Inductors . . . 912
25.13.4 Virtues of Grounding . . . 913
25.13.5 Passive Mixing . . . 914
25.13.6 Devolved Mixing . . . 914
25.13.7 Balanced Mixing . . . 915
25.13.8 Pan Pots . . . 915
25.13.9 Surround Panning . . . 916
25.14 Monitoring . . . 917
25.14.1 Solo, Solo-Free, and Prefade Listen . . . 918
25.14.2 Monitoring Controls . . . 918
25.14.3 Related Crosstalk . . . 919
25.14.4 Unrelated Crosstalk . . . 919
25.14.5 Quantifying Crosstalk . . . 919
25.14.6 Meters . . . 919
25.14.6.1 VU Meters . . . 920
25.14.6.2 Peak Program Meter (PPM) . . . 920
25.14.6.3 Other Metering . . . 922
25.15 A Typical Multitrack Console Described . . . 922
25.15.1 Channel System Example . . . 922
25.15.2 Function Modes . . . 923
25.15.3 Recording Mode . . . 923
25.15.4 Mixdown Mode . . . 923
25.15.5 Overdub Mode . . . 923
25.15.6 Logic Control . . . 923
25.15.7 Switch Debouncing . . . 926
25.15.8 Logic Sense . . . 928
25.15.9 What Is a PROM? . . . 928
25.15.10 Logic Meets Analog . . . 929
25.15.11 Auxiliary Channel Feeds . . . 930
25.15.12 Summing Modules . . . 930
25.15.13 Mixes to Outputs . . . 932
25.15.14 Master Functions . . . 932
25.15.15 Power on Reset . . . 933
25.15.16 Meters and Head room . . . 933
25.15.17 System Level Architecture . . . 933
25.16 Consoles and Computers . . . 935
25.16.1 Fader Automation . . . 936
25.16.2 VCAs . . . 936
25.16.2.1 Transistor Junctions . . . 937
25.16.2.2 Gilbert Cell . . . 937
25.16.2.3 Log-Antilog Cell . . . 938
25.16.2.4 Commercial VCAs . . . 938
25.16.2.5 Control Voltage Noise . . . 938
25.16.2.6 VCA Channel Application . . . 939
25.16.3 Digitally Controlled Amplifiers . . . 940
25.16.3.1 Multiplying DACs . . . 940
25.16.3.2 Charge Injection . . . 940
25.16.3.3 Zero-Crossing . . . 941
25.16.4 Discrete Logic and Programmable Gate Arrays . . . 942
25.16.5 Recall and Reset . . . 942

25.16.6 Data Acquisition 942
25.16.7 Control A/D Conversion: Central or Distributed? 945
25.16.8 Recall Display 945
25.16.9 Nulling. 946
25.16.10 Resetting Functions 946
25.16.11 Motorized Pots and Faders 946
25.16.12 Resolvers 947
25.16.13 Control Surfaces 948
25.16.13.1 The Single Channel Concept 948
25.16.13.2 Commercial Console Control Surfaces 949
25.16.13.3 Control Surface Intelligence 951
25.16.13.4 Multiuser, Multifunction 952
25.16.14 Goodbye Jackfields, Hello Routing. 952
25.16.15 Integrated Control of Outboard Gear 953
25.16.16 Digital Audio Workstations (DAWs) 953
25.17 Digital Consoles 953
25.17.1 Digital Audio Systems. 954
25.18 converters 954
25.18.1 A/D Conversion. 954
25.18.1.1 Resolution 954
25.18.1.2 Sampling Rate 955
25.18.1.3 Convertor Limitations and Requirements 955
25.18.1.4 AntiAliasing Filters 956
25.18.1.5 Types of A/D converters 956
25.18.2 Digital-to-Analog Conversion (DAC). 957
25.18.2.1 Conventional Ladder DACs 957
25.18.2.2 Reconstruction Filters 958
25.18.2.3 Oversampling 958
25.18.2.4 Sigma-Delta DACs 958
25.18.3 Sample-Rate converters (SRCs) 958
25.19 Digital Signal Processors 959
25.19.1 Multiplier/Accumulator (MAC) 959
25.19.2 Instruction Cycles 959
25.19.3 Processor Types. 960
25.19.4 Floating Point 960
25.19.5 Parallelism 961
25.19.6 Multiple Memory Spaces 961
25.19.7 Real-Time Specific Peripherals. 961
25.20 Time Processing in DSP 961
25.20.1 Time Manipulation 961
25.20.2 Delay 961
25.20.3 Echo, Echo, Echo 962
25.20.4 Reverberation 962
25.20.5 Averaging 962
25.21 DSP Filtering and Equalization 963
25.21.1 Transversal, Blumlein, or FIR Filters 963
25.21.1.1 Impulse Response 963
25.21.1.2 Windowing 964
25.21.1.3 Symmetrical FIRs 965
25.21.2 Recursive Processing. 965

    - 25.21.2.1 IIR Filters . . . 966
    - 25.21.2.2 The Biquad . . . 966
    - 25.21.2.3 Coefficient Analysis . . . 967
    - 25.21.2.4 Filter Quantization Distortion . . . 967
    - 25.21.2.5 Cascaded Biquads . . . 968
  - 25.21.3 Parametric EQ . . . 968
  - 25.21.4 Shelving EQ . . . 969
  - 25.21.5 EQ High-Frequency Response Anomalies . . . 970
- 25.22 Digital Dynamics . . . 970
  - 25.22.1 Mapping Dynamics . . . 971
  - 25.22.2 Literal Dynamics . . . 972
    - 25.22.2.1 A Simple Digital Limiter . . . 972
    - 25.22.2.2 Feedback-Style Limiting and Compression . . . 973
    - 25.22.2.3 Gating . . . 973
  - 25.22.3 Inadequate Samples for Limiters . . . 974
  - 25.22.4 PreDelays, or Look-Ahead . . . 975
- 25.23 Digital Mixer Architectures . . . 975
  - 25.23.1 The Sea of DSPs approach (sometimes known as a Sharc Tank) . . . 975
  - 25.23.2 A Practical Digital Mixer . . . 976
    - 25.23.2.1 Serial Audio Format . . . 976
    - 25.23.2.2 Inputs . . . 976
    - 25.23.2.3 Mix Stage . . . 976
    - 25.23.2.4 Mix Code . . . 977
    - 25.23.2.5 Universal Mix Buses . . . 978
    - 25.23.2.6 Coefficient Compounding . . . 979
    - 25.23.2.7 Coefficient Slewing . . . 979
    - 25.23.2.8 Clocking . . . 979
  - 25.23.3 Signal-Processing Control . . . 980
    - 25.23.3.1 Controlling the DSPs . . . 980
    - 25.23.3.2 Metering . . . 980
    - 25.23.3.3 Host Microcontroller . . . 981
- 25.24 Digital Audio Workstations (DAWs) . . . 981
  - 25.24.1 The PC . . . 983
  - 25.24.2 MIDI Sequencing—Where It Began . . . 983
  - 25.24.3 DAW Audio Ins and Outs . . . 984
  - 25.24.4 DAW Internal Audio Paths . . . 984
  - 25.24.5 Plug-Ins . . . 984
  - 25.24.6 DAW Limitations . . . 985
- 25.25 Moving Digital Audio Around . . . 985
  - 25.25.1 Moving Audio—small-scale . . . 986
    - 25.25.1.1 AES-3 . . . 986
    - 25.25.1.2 AES-42 . . . 986
  - 25.25.2 Moving Audio—Multiple Paths . . . 987
    - 25.25.2.1 AES-10—MADI . . . 987
    - 25.25.2.2 ADAT . . . 988
    - 25.25.2.3 USB . . . 988
  - 25.25.3 Digital Audio Networking . . . 989
    - 25.25.3.1 CAT-5/RJ-45 interconnectivity types . . . 989
    - 25.25.3.2 TCP/IP—Audio over Ethernet . . . 989
- 25.25.3.2.1 Buffering Latency . . . 990

25.25.3.2.2 Latency—How Much Is Too Much? 991
    25.25.3.3 UDP 991
25.25.3.3.1 Serial or Loop Networking. 992
25.25.3.3.2 Star Network Topology. 992
25.25.3.3.3 Mixing in the Router 993

This page intentionally left blank

25.1 Introduction

Mixing consoles are an immense subject; their understanding though is a key to professional audio. In this section consoles are discussed from the basics of their architectures, features, design elements, and all the way through to DSP and its implementation in digital mixer design. Consoles abound in forms quite unlike the traditional big sea of knobs. Productions huge and modest are regularly done with a screen and a mouse—indeed, nearly everywhere in operations that don't require immediate access to controls, mostly live. But consoles they are; they're just hiding in unfamiliar shells; their schema are, despite outward appearances, directly traceable to traditional audio architectures.

Commercial mixing consoles live or die not just on how closely they fit their particular application; a several-hundred-thousand-dollar buy decision is often made on things as fiercely disparate as more or less favorable finance schemes and the way consoles sound (or more often are reputed to sound). This section does not include console cost other than drawing distinctions between straightforward and extravagant approaches, but it does explore many aspects of what makes consoles sound better or worse.

Along the way, explanations are given of what each common control does and how it is normally used and why. Examples of how they have developed and been implemented in electronics are also given. A range of console arrangements (architectures) gives enough clues to analyze how any encountered system actually operates. Description of circuitry and techniques is less theoretically driven than practically derived, with no apologies given for blow-by-blow analyses of real commercial multitrack mixing console designs. It is hoped that this will augment and lend perspective to earlier descriptions of typical circuit blocks.

Seemingly the only thing preventing most mixers from being digital nowadays is that the cost benefit for many applications has not yet tilted far enough in that direction; although hardly mature, the technology and the available quality are not impediments. Given that, one might question why this chapter retains a lot of "analog stuff." The answers are manifold: A lot, arguably still most, of mixers in use today are still analog and far from a prediction made in the mid-eighties that the Last Great Analog Console had probably already been built, manufacturers both established and new seem to think it worthwhile to wheel out new analog behemoths once in a while. And, almost in mirror fashion, at the low end of the market economies of scale and slim margins still preclude digital on cost alone, where any meaningful control surface is needed. But those aren't the main reasons consoles exploded, evolved, and matured operationally in the same era as similarly burgeoning analog technology; the technology inevitably influenced the application with its own rationalized costs and limitations. The applications—consoles and their constituent signal-processing elements—are nowadays being emulated digitally. Yes, a huge amount of the engineering of digital mixing consoles is in accurately recreating foibles inherited from their analog ancestors, good or ill. It behooves one to understand why things ended up the way they did, so that one can optimally progress into the new domain. It's called learning from history, only in this case the history is still very much alive and has its teeth.

An overview of digital signal processing as applied to consoles will, as deeply as it is possible to go before nasty equations arise, give an insight into how these things work—at first blush it all seems black magic—really it's just a lot of black chips each with a specific and usually straightforward purpose. Similar to the way that a real analog console design is dissected and explained in the following pages, a real digital console design is broken down for overview and analysis.

Parallel, and indeed prior, to the incursion of DSP audio was digital control of audio. Although DSP and digital control necessarily go hand in hand, digital control of analog techniques are overviewed here, too.

25.1.1 Console Development

The establishment of consoles was a slow and gradual process. Similarly, systems—or preorganized arrangements of devices—evolved slowly, too. In most audio work the two are now considered as almost synonymous; the greatest departure from this is the inclusion of a console as part of a system. But even then, there is no doubting that the console is the heart and substance of the system.

The history of consoles reaches back to the time when the recording process was purely mechanical, Fig. 25-1, followed by its electrical analog, Fig. 25-2, which included a source transducer (in this instance a microphone), a means of gain (an amplifier), and an output transducer (a disk-cutting head). It doesn't take a staggering amount of imagination to extend this system to embrace other applications: public address, acoustic enhancement of natural sound by electronic means, Fig. 25-3; disk replay, Fig. 25-4; and broadcasting by replacement of a simple electromechanical transducer by a radio transmitter, Fig. 25-5. The objective of the system is to facilitate the transfer of a signal from one source—be it a simple transducer or another system—to a destination.

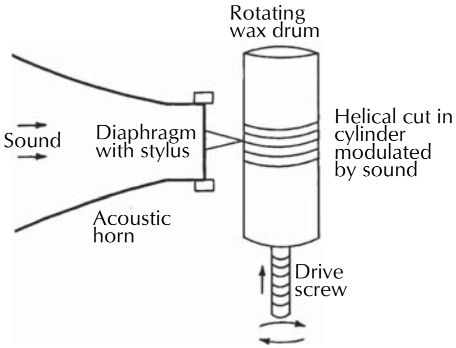


Figure 25-1. Mechanical recording or early drum recording.

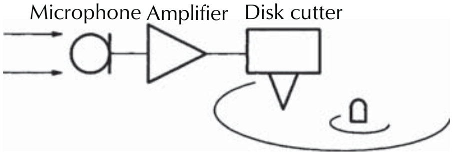


Figure 25-2. Electrical recording (disk cutter driven by electricity).

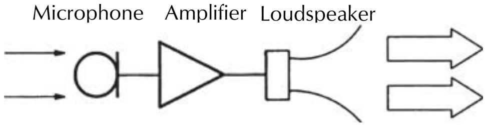


Figure 25-3. Public address system.

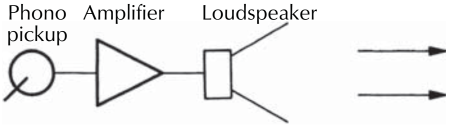


Figure 25-4. Disc replay.

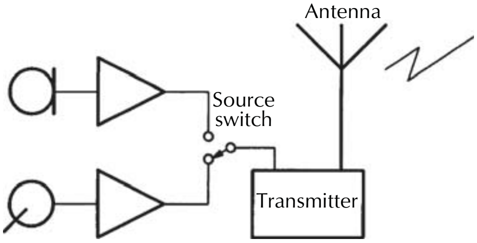


Figure 25-5. Simplified broadcast system.

Of course things get a bit more involved than that, and to demonstrate this complexity, the evolution of what is probably the most important subsystem to our industry—the recorder, mono, stereo, or multitrack—will be used to explain how it, almost single-handedly, made everything as complicated as it is today. Disks were permanent. You got it right or you didn't. Tape at least gave the chance of one more take.

Mixing in the early days of system development was surprisingly easily achieved—just connecting the outputs of the various source input amplifiers together did it perfectly adequately. It's important to understand that the technology of the day facilitated this simplicity far more, paradoxically, than today's gear. Tube amplifiers, such as were then used, needed to be terminated at their outputs by a specific impedance for proper operation, which for reasons discussed later was universally 600 Ω floating balanced. By simply connecting amplifier outputs together, a mix of sources was achieved, provided each of the source amplifiers saw 600 Ω. It was only a very minor step for interspersed networks to become constant-impedance variable attenuators, usually in the form of rotary controls. The *pot* (from potentiometer) or *fader* was born. The ability to create a balance of sundry sources for the chosen destination is perhaps the most recognized feature of the console and its system. Convention and common sense rule this as the main signal path, and other paths are subsidiary or auxiliary to it.

25.2 Auxiliary Paths

25.2.1 Monitoring

Take the example of Fig. 25-6, where a single microphone is being laid on a recorder. It's operationally necessary for the system operator to hear the signal going to the recorder with headphones or a control-room monitor loudspeaker. To facilitate this requirement, a parallel feed is taken off the machine input to the operator's monitor. *Monitoring* is perhaps the most important of the auxiliary signal paths; upon it is based the qualitative decisions of the nature of the signal in the main path. It is the reference.

Fig. 25-7 applies a small extension to the basic monitoring path in the form of a source/replay switch, enabling operators to hear the aftermath of their efforts. If the recorder has separate record and play signal paths, they can even toggle between the two while actually recording for immediate quality assessment. The monitoring section is born.

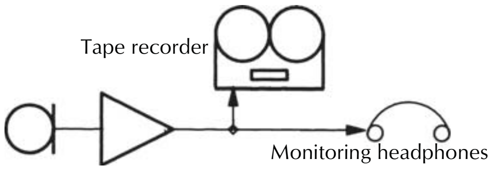


Figure 25-6. Simple microphone-to-recorder monitoring (source only).

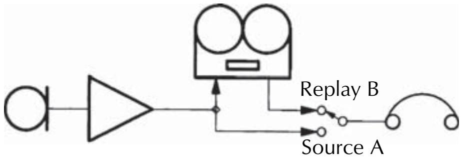


Figure 25-7. Recorder monitoring (source and tape).

25.2.2 Prefade Listen and Audition

When a multiple-source system is established (similar to Fig. 25-8), another monitoring requirement, *prefade listen* (PFL) and *audition*, is required. Imagine the case of a radio broadcaster, where the sources consist of disk replay units and microphones; it's an obvious necessity to be able to listen to a source prior to its being put on air to check that:

1. The microphone is set at the correct position, level, or even working!
2. The required section of a disk or tape is cued up or ready to play.

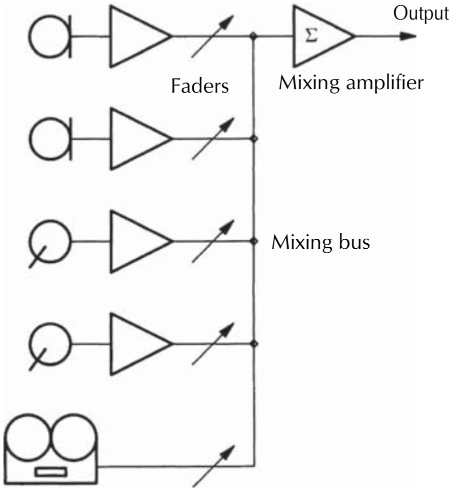
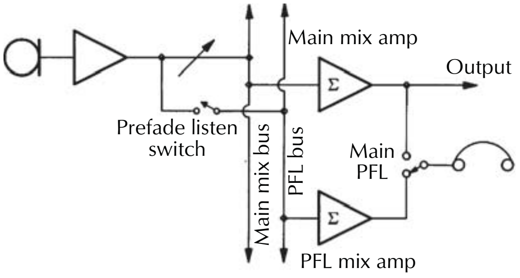
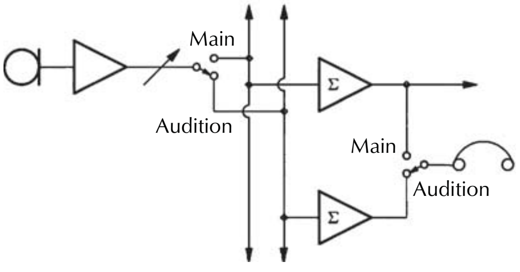


Figure 25-8. Multisource mixer.

There are two basic methods of arranging this prehear function, as shown in Fig. 25-9. They owe their existence primarily to slightly different operating practices on opposite sides of the Atlantic. The first, Fig. 25-9A, involves switching the signal immediately prior to the fader on the selected source path into the monitoring chain. This is called *prefade listen (*PFL*)*. A useful but not immediately obvious virtue of this arrangement is that it is possible to listen to a channel's contribution to a mix of which it is part without disturbing that mix. It is, therefore, a nondestructive monitoring function. The alternative method shown in Fig. 25-9B consists of removing the required channel (postfader) from the mix and placing it onto a second parallel mix facility, commonly called *audition* or *rehearse*; it is possible in this mix to emulate exactly what would happen in the real mix without upsetting the presently active mix. A disadvantage of this method is the inability to use the function when the channel is live because it disrupts that source and prevents it from going to the mix. It is a destructive monitoring technique. Each method has its virtues, though, and most modern consoles use both techniques to varying extent. That said, all but really small American broadcast consoles nowadays employ a PFL-type cue function, the audition/rehearse bus being arranged to be a secondary mix bus independent of but in the same vein as the program bus. The name often remains as an echo of its original function. Postfader monitoring, however, lives on in large production consoles as *in-place stereo* monitoring, described later.



A. Prefade listen.



B. Audition.

Figure 25-9. Prefade listen (PFL) and audition.

25.2.3 Overdubbing and Foldback (Cue)

While broadcasting lends itself to explaining the need for individual channel monitoring, original material generation onto tape serves best to explain another crucial auxiliary signal path.

It didn't take long before studios were using more than one tape machine in a technique known as *overdubbing*. Briefly, this involved recording a backing track (for instance a rhythm section) on one machine and then playing that back while vocalists sang or soloists played along with it; the whole was mixed together and recorded on a second machine (bounced) as shown in Fig. 25-10. This could be carried on until the subsequent machine-to-machine generation losses became too objectionable (although that never seemed to bother many early producers!). (*Generational losses* occur because the tape machines of the era were less than perfect, what came out of them being noticeably ropier than what went in!) Naturally, it was essential that the musicians in the studio were able to hear via loudspeakers or headphones that to which they were supposedly playing along; this is where *foldback* (*cue*) comes in. In its simplest form, it could be a straight derivative of the main mix output, since this output has basically everything necessary in it. This system, however, has a few shortcomings due primarily to conflicts between what the final mix is intended to be and what the artist(s) needs to hear to perform satisfactorily. A prime example of this dilemma is in the recording of the backup vocalists sections; usually they take a fairly minor part in a mix, being balanced well down. Contrary to this is the need of the vocalists to not only hear the track played back to them but to hear themselves sufficiently well—usually enhanced—to pitch and phrase themselves effectively. These conditions are next to impossible at the final mix. A solution lies in Fig. 25-11 where a separate balance of the relevant sources is taken and fed separately to the performers, giving them what they most need, a foldback mix. The takeoff for the foldback feeds is almost invariably prefader so that the artist's balance remains unaffected regardless of what modifications may be necessary for the main mix.

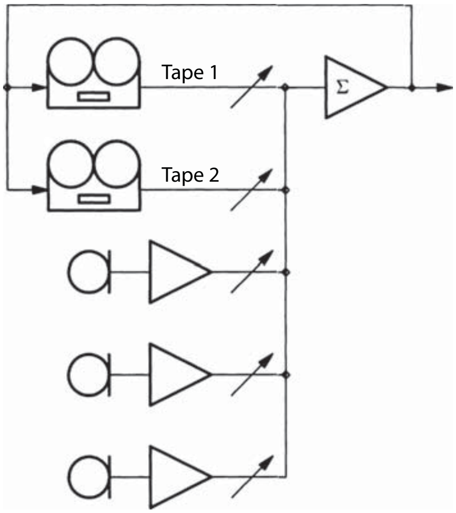


Figure 25-10. Overdubbing/bouncing, a previous microphone mixer recorded on tape 1 may be played back along with a further microphone mixer onto tape 2 and vice versa.

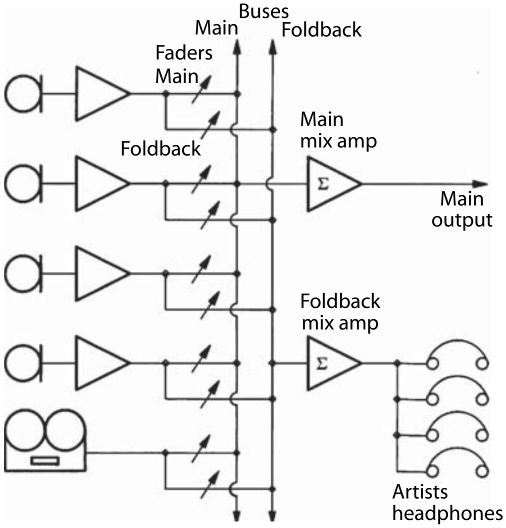


Figure 25-11. Foldback mix.

25.2.4 Echo, Reverberation, and Effects Send

The move (regrettable as it may seem) from natural performing acoustic environments to the more cultured, drier, closer miced techniques brought with it many problems attendant to the advantages. How do you make a sound seem as though it was recorded in a great concert hall if it was done in a small studio? Reverberant chambers were an initial answer, being relatively small rooms acoustically treated to have an extended reverberation time (bathroom effect). Driven obliquely at one end or corner by a loudspeaker(s) and sensed by a microphone(s) at the other end, which is amplified and balanced into to the main mix, a fairly convincing large room reverberant effect can be achieved. Simplistically, all that's needed to feed the loudspeaker in this room is a derivation of the main mix, but similar to the problems with foldback mixes, artistic judgments dictate something more complex. Some instruments and sounds benefit greatly from being dry (most of a drum kit, for example), while others—vocals, in particular—sound quite dry, cold, and uninteresting. A means of adjusting the relative amounts of artificial reverberation due to various sources would be beneficial. Fig. 25-12 shows a small console system complete with an echo send mix bus (echo in this sense including reverberation); the echo return is brought back into the main mix just as any additional source would be. Echo feeds are nearly always taken postfader, keeping the reverberation content directly proportional (once set) to the corresponding dry signal in the mix regardless of the main channel fader setting.

Today, any number of foldbacks and effects sends are in use as toys (effects boxes) proliferate; long gone are the days when one of each feed was sufficient.

25.2.5 Communications (Talkback)

An often mentally mislaid but crucial console auxiliary path is *talkback*; that is, the ability of the console operator/producer to talk to various people involved in the recording. The primary need for talkback is to be able to communicate with the studio area that is necessarily acoustically separate from the control/monitoring room. Since there are already foldback feeds going to the studio area for performer cues, it makes sense to talk down these feeds, which is talk to foldback (talk to studio). Another useful function in this vein is *slate*. This curiously named facility allows the operator to talk into the main mix output and thus onto tape for track and take identification purposes.

25.2.6 Combined Auxiliaries

In summary, a usable console has to have several signal paths in addition to the main mix path. These include overall and presource monitoring, prefader adjustable foldback feeds, postfader artificial reverberation feeds, and communication (talkback) feeds as shown in Fig. 25-12.

25.3 Stereo Consoles

Stereo predated multitrack recording. Technically the required console techniques were not very far removed from those just described. Assuming the same bouncing (machine-to-machine overlay and transfer system described earlier in the section on overdubbing), stereo just means two of everything in the main signal path.

25.3.1 Panning

Panning is the technique of positioning a single monophonic source within a stereophonic image. It isn't true stereo; true stereo can only be achieved from coincidentally aligned microphones. Instead, it is panned mono. Simply, the ear is deceived by pure level differences between the left and right paths of a stereo pair into perceiving differing image position; fortunately for the entire industry, this is a trick that works rather well and is quite simply realized.

Complementary attenuators (one increasing and one reducing attenuation, with rotation) feeding the L and R mix paths from a mono source is the most common method. Fig. 25-13A illustrates this system. The pan pot is usually inserted after the source fader. An alternative arrangement is shown in Fig. 25-13B. Here the pan pot is inserted prior to the fader; a ganged matched fader is required with this method. This arrangement can be useful when stereo PFL is required, although there are other ways of achieving stereo in-place monitoring for sources that will be described later.

25.3.2 Auxiliaries for Stereo

Auxiliary paths remain largely untouched by the upgrade to stereo of the main mix path; the monitoring section stays just the same in systemic function (but obviously with two paths instead of one to cope with stereo feeds). Both the prefader foldback and PFL take-offs are still in mono. The postfader echo-send feed is usually taken out before the main path pan pot, so they remain mono, but the returns pass through their own pan pots such that the reverberant image may also be spatially determined in the mix. It's become normal practice to make echo-send feeds stereo in their own right, Fig. 25-14, via their own pan pots' mixing to two outputs. Many reverberation rooms, plates, and boxes are capable of supporting a diffuse stereo field. The purpose of this is to excite the reverberant chamber (or plate or springs or little black box) spatially, conjuring a more solid and credible reverberative effect in the main mix. If a panned echo-send output isn't available, it's common to use a pair of separate postfader feeds and juggle the levels between them.

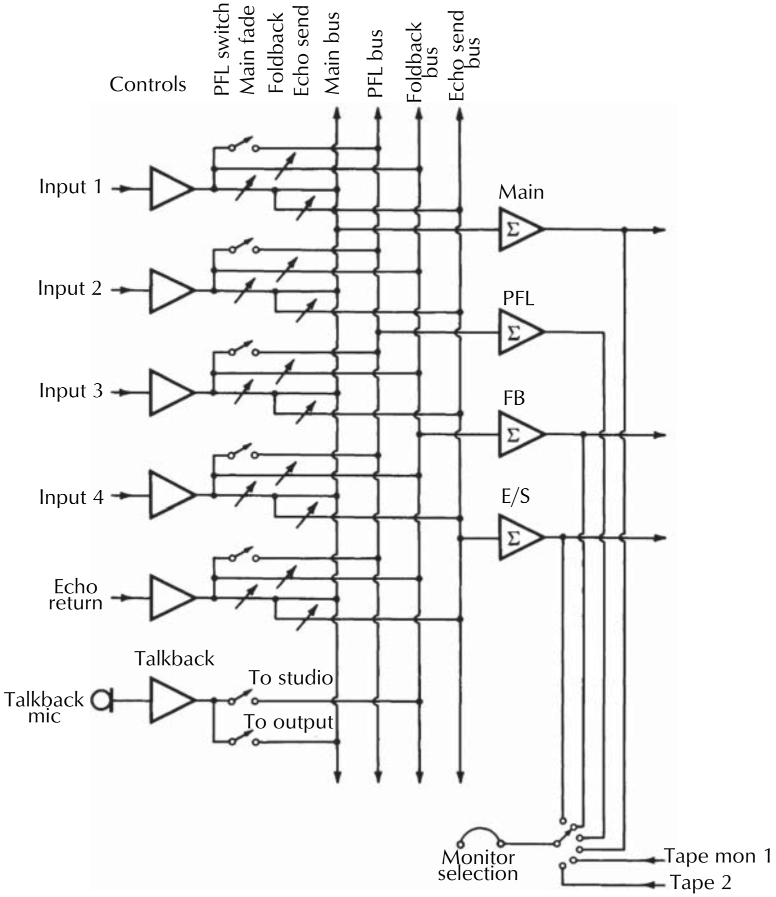

Controls
PFL switch
Main fade
Foldback
Echo send
Main bus
PFL bus
Foldback bus
Echo send bus
Input 1
Input 2
Input 3
Input 4
Echo return
Talkback
Talkback mic
To studio
To output
Main
PFL
FB
E/S
Σ
Monitor selection
Tape mon 1
Tape 2

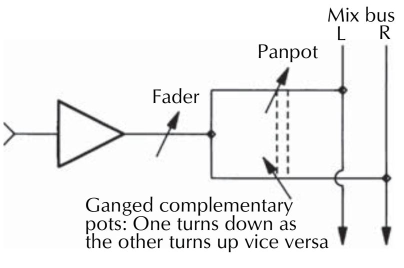
Figure 25-12. Small mixer system showing auxiliary functions.

25.3.3 Multiple Effect Feeds

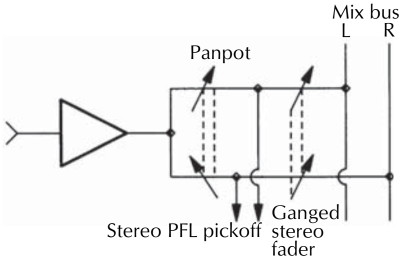
Currently, there is a whole gamut of electronic toys applied to mixdown to achieve specific sounds: Harmonizers®, delays, flangers, phasers, automatic panners, artificial reverberators of various sorts, and so on. These all need to be fed from their own effects mix paths. Similarly, studio foldback mixes have grown more profuse with changing music and increasing musician sophistication and awareness of studio techniques; consequently, the number of auxiliary mixes within modern consoles has risen quickly. A rationalization of this is to make those auxiliary mixes multipurpose, usually by allowing them to be switchable between prefader and postfader feeds on their appropriate sources. A smaller number of buses are needed in this way; during the recording process the emphasis is on many foldback (prefade) feeds for the musicians in the studio and maybe one or two toys to spice up the monitoring. On the other hand, during the overdubbing and mixdown phases very few foldbacks (if any) are needed but every bus will be set to postfade and laden with effects. Additionally, in broadcast it is commonly necessary to talk back down some or all of these mixes individually; such a feed is called an *interruptible foldback*, or IFB. Large modern consoles for live applications often have many stereo foldback auxiliaries, driven by the trend toward the (usually wireless) in-ear headphones beloved by performers.

25.4 Dawning of Multitrack

Multitrack operation is when a number of separate parts of the recording are laid onto separate tracks on a recording machine and subsequently remixed down onto another machine (be it mono or stereo) or even interim bounced onto spare tracks on the same machine.


Mix bus
L
R
Panpot
Fader
Ganged complementary pots: One turns down as the other turns up vice versa

A. Postfade.



B. Prefade.

Figure 25-13. Panpots.

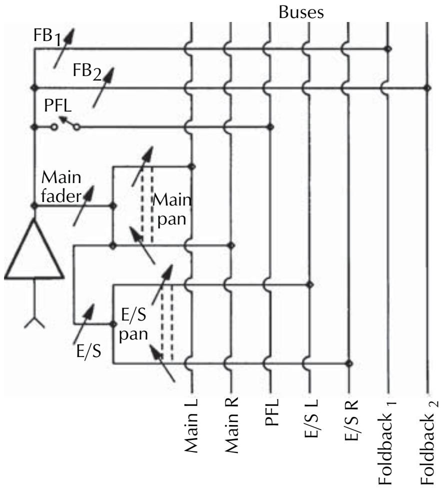


Figure 25-14. Channel feeds showing foldback and stereo echo-send feeds.

Stereo recording using two-track tape technology seemed to many to be the zenith of professional audio. Many will argue the point even today. There is inescapable evidence of the validity of that opinion in that some of the finest stereo recordings, especially of classical and jazz works, were done using fundamental microphone techniques straight onto two track. Even in the field of pop records where things were bounced mercilessly the final master still represents the first generation of the last overdub. (In retrospect that is an advantage over contemporary multitracking where the master is at best the second generation of everything.)

Multitrack soon reared its head(s?) in the early 1960s—initially as three track and four track across 1 inch tape; there are those who regard that as the zenith. More tracks reduced the number of intermachine bounces, but they still added up! *Sergeant Pepper* is of the bounced four-track genre (although many parts were done on a pair of loosely synced machines for pseudo eight-track) and stands up rather well even today. It does put things in perspective. How much more technology is needed for what?

Three tracks afforded a great advantage over two tracks for modern music producers at the time. Two-track recordings were always hampered by the need to make sure that all the earlier things done in a bouncing sequence were right to begin with; there was no chance of subsequently altering them. Three-track recordings, typically in a Track/Vocals/The Rest format, took a little of that pressure away. Already producers and performers were taking advantage of the multilayered production approach to take the heat out of recording; it was no longer necessary for everyone from lead vocalist to third trianglist to be present all at once for a momentous occasion. Bits could be done one at a time. The extension to this given by multitrack is simple to see: the more tracks, the smaller those bits need be and the fewer things needed to be incontrovertibly mixed. Putting off the day of reckoning—the final mixdown—is one of the strongest appeals of multitrack. This, indeed, has led to a curious polarization in the business; *tracking*, the laying down of individual tracks, is typically done in entirely different studios or environments to *mixing*. And *remixing*, the construction of yet different mixes from the same basic tracks for specific genres such as dance mixes, has spun off into yet another subindustry. So much for making spontaneous music.

25.5 Grouping and the Monitoring Section

Each signal source in the console needs some routing to determine the machine track on which it is going to end up. It's a situation that hardly existed previously, since it was pretty sure that the mono or stereo output of the console was going to go straight to the respective mono/stereo inputs on the tape machine(s). There were on a stereo console just two groups where all the sources were summed together; for multitrack as many groups as there are tape tracks are switch selectable from the sources—any source to any machine track. The alternative hard way, patching everything across on a jackfield, was and is exceedingly tedious, messy, expensive, and error prone.

Four-track recording set the mold for console design for many years. The monitoring section evolved. Fig. 25-15 can be compared to the simpler back end of a stereo mixer in Fig. 25-12. The main difference can be seen as the addition of an entirely separate mixer within the console just to handle the multitrack monitoring. Fortunately, it's a fairly bare-bones mixer; it's all at high signal levels, and little, if any, gain is required except as makeup gain in the monitor mix bus.

While all these tracks are being laid, it's necessary to hear what has been done previously in the control room and studio. In the same way that source/return listen of stereo machines was needed, so each individual track of a multitrack needed similar treatment. It grew, though. Initially, as the number of tracks per machine increased, the number of mixer groups increased correspondingly. Each group had its own A/B switch relating to that individual console track output and the associated machine return, with its own level and pan controls feeding an altogether separate stereo monitor mix. This new monitor mix appeared as another source on the main monitor selector. This, alas, was insufficient. Foldback prefade mix feeds no longer became a luxury but a necessity, since the desk stereo output or a derivation thereof could no longer be relied upon to be even roughly what the artist needed to hear. There was no proper console stereo output at any time other than mixdown. Foldback feeds were added to the monitor system on each group. Effect sends were also added, just to help the monitoring sound pretty.

The monster has split itself amoebalike into two entirely separate signal-processing systems: the main mixer and a monitor mixer. A curious situation occurs: the mix used for monitoring during the original multitrack recording had to be transferred over to the main system entirely at some time for mixdown. Ordinarily, tape-machine returns are not only brought back into the monitoring section but are also tied to high-level line inputs on the main mixer section. The remix takes place using those channels into the main stereo mix bus.

Perhaps the first major rationalization (which occurred long after many conventional X-input, 24-group, 24-monitoring consoles had been made) was a result of the realization that few people actually needed 24-group faders sitting there full up, collecting dust. Losing them instantly avoids a normally unnecessary gain-variable stage in the signal path, which, if maladjusted, could upset noise or headroom performance.

Individual channel outputs together with a much smaller number of stereo mixing subgroups—usually four or eight pairs—which could again be routed to any of the multitracks, proved easily as flexible. But still there was duplication of monitor buses and main stereo mixing buses both with their attendant effects and foldback feeds rarely being used simultaneously. At last the dawning of the realization that the pair, that is, the monitoring and stereo mastering buses, could be one and the same thing. In-line monitoring recording

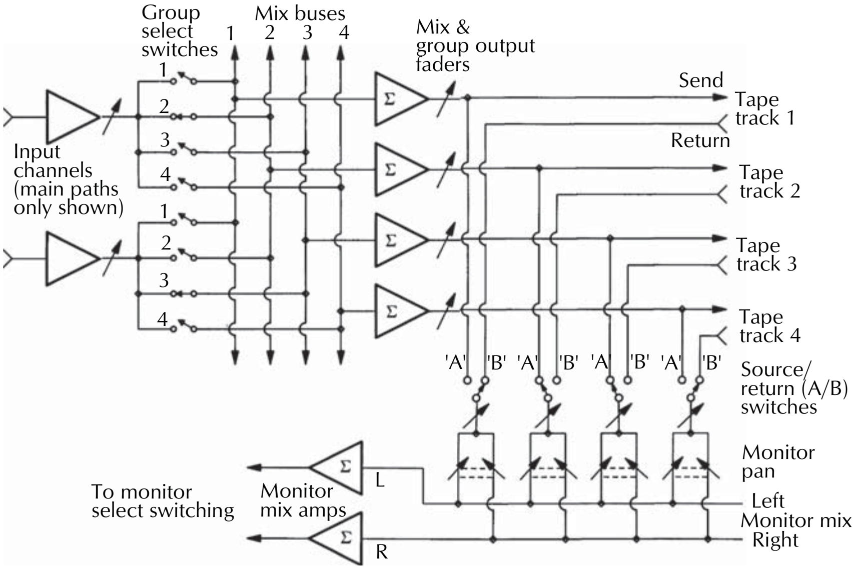


Figure 25-15. Four-track monitoring.

systems had come to fitful fruition. The in-line console includes all of a recording channel's processing and all of a machine return's monitoring controls within one channel strip; it allows efficient sharing of controls, processing, and mixes between those paths, maximizing their full utilization through the tracking, overdubbing, and mixing phases of a production. It also does away with the separate multitrack monitoring section, the mixer within the mixer that nearly doubled the physical width of conventional consoles.

We all have to be thankful for the cranks and visionaries along the way (often the same) who have manipulated or shocked the industry into grudgingly lurching back into step with technology's capabilities. These developmental milestones represent significant plateaus of thinking that form the basis of today's console concepts. In-line is a classic example.

25.5.1 Subgrouping and Output Matrices

Particularly in live applications (e.g., sound reinforcement or broadcasting) the ability to make a subgroup of related sources—say drum mics, bass, guitar, keys, backing vocals, lead vocal (each of which can have many sources themselves)—and then rebalance them together is a valuable addition. (This means that instead of having to gingerly pull down the 10 mics on a kit without destroying the previously hard-won balance, a single fader on that subgroup can be moved instead.) These are real subgroups, so called because a real mix of real audio sources is created, rather than a similar overall result happening by way of a VCA subgroup (described fully later) in which only the fader movements are tied. An output is available just containing the subgroup member sources, useful if processing (EQ, dynamics, etc.) is required over them exclusive to other sources such as auxiliary sends for the addition of effects solely to the subgroup and remixing. This latter is a particularly powerful use for these subgroups; feeding them as sources into a downstream mixer, often called a *matrix mixer*, from which an often large number of matrix output mixes are created.

Again using sound reinforcement as an operational example, the many performers on stage all need to hear both themselves and the rest of the performers either in monitor speakers or in personal earpieces; the trouble is, the balance that each of these people needs is typically entirely different! Using the individual remix capability on each matrix output fed by the earlier-created subgroups, many different mixes of the same few subgroups are possible, hopefully resulting in a calm stage.

25.6 Console Design Developments

Two distinct considerations interplay in determining the ability of a console to fulfill a given application. These two—the system and the electronics—have entirely differing parameters that need to be defined but are, nevertheless, completely indivisible.

The electronics, as much as being designed to perform required functions, must be very carefully designed not to be a major influence on the sound of the console. Most causes of sonic disturbance can be attributed or predicted, and still dubious circuit configurations can be avoided altogether. There seems to be a groundswell of designing sonic character back into studio electronics; this after generations of striving for accuracy and neutrality is a touch alarming. The good news is that consoles (unless otherwise eccentrically contrived) are still expected to be neutral, the color being acquired by the gallon in external rack boxes. To that end, unless specifically stated, the electronics described here are intended to be neutral sounding. To the shock of some purists, commonly available integrated circuit operational amplifiers are generally used throughout the designs in this chapter. The reasons why (other than the obvious convenience), together with the reasons why they acquired a bad reputation, are treated in depth in Section 25.7.

Operational amplifiers (op-amps) have, in recent years, revolutionized the concepts and systems capability of full-performance audio consoles. Their use allows system elements to be thought of, designed, and implemented as building blocks. This simplifies matters considerably, but it also entertains the valid criticism that console design can be relegated to a do-it-by-numbers routine. Fortunately, device idiosyncrasies, subtleties, and the entirely separate science of getting heaps of individual system elements to behave successfully as a total console prevent this.

Fortunately for the console industry, the large proportion of the current console manufacturers started off in life as small groups of musicians and studio engineers furtively constructing mixers for their own ends, resulting in grass-roots system design owing everything to immediate operational needs. Continuing in this vein in production, the manufacturers are listening to and, most importantly, relating to customer needs because they've played this game for themselves.

Not too long ago, systems and mixers as such didn't exist. All the bits of electronics used in the control room sat there with all their inputs and outputs accessible by way of a jackfield for the prosperous or by small screwdriver and sore knees for those who weren't.

Mixing sources was accomplished by directly paralleling amplifier outputs (possible because all the old tube gear was designed with a particular termination impedance in mind, usually arranged to be a conventional balanced 600 Ω) and either hoping or arranging that the destination had enough gain to make up accrued paralleling losses. Crude as that may seem today from an engineering viewpoint, it has a sheen of pure elegance. An amplifier was just that, a box that had a balanced 600 Ω source and termination impedance. It might also have an alternative bridging (> 10 kΩ) input terminal and a selectable amount of gain offering universal application from microphone amplifiers through mixing amplifiers to headphone amplifiers. To do more things, more boxes were added. Equalizers and limiters, a treasured few if there were any, were similarly universally applicable. Variable-level control was again attained by true balanced 600 Ω source and termination, via studded rotary attenuators. The utter beauty of the systemless studio was that anything could go to anywhere via anything else and be mixed or distributed at any point on the way.

Soon enough amplifiers were hardwired to attenuators and designated specifically a microphone amplifier, and a system had been created. Some of these together with a mixing gain makeup amplifier were thrown in a box. The mixer was born.

It has been downhill ever since, with ever-increasing numbers of system elements being tied together in increasingly knotted manners in order to maintain some kind of flexibility. Perversely, a system can be defined as a means of reducing the ultimate versatility of its constituent parts.

Once a mixer was accepted as a system element itself, the problem set in further. There was no need to provide for convenient connection of its internal interconnections to the outside world, so the balancing transformers disappeared, and more economic alternatives to the stud attenuators operating at more convenient internal impedances evolved. By a more positive token, the electronics were gradually becoming optimized for the specific functions to which they were designated, such as the microphone amplifier and the mixing amplifier. (The question nags us whether a universal amplifier, by now all but obsolete, could be optimized for all the varying requirements, this is unlikely.) Still, at least all the inputs and outputs of the mixer were conventional. This held true until the slow demise of vacuum tubes in professional audio.

25.6.1 Transistors

Transistors were justifiably unpopular for a long time because of the numerous limitations they placed on design. The headroom was severely limited because of the low supply voltages that could be applied to the early devices. They were noisy. The lower operating impedances and differing modes to tubes took some getting used to and, when they clipped, they actually clipped rather than gracefully bending (characteristic of tubes that people had known, loved, and frequently taken advantage of even now). To realize a reasonably low stage distortion, many transistors in compound configurations using heavy amounts of negative feedback were used—a far cry from a single tube stage operating virtually wide open with little feedback. This gave rise to a peculiar phenomenon that sounded as if it hailed from science fiction—*zero impedance*.

The heavy negative voltage feedback employed around transistor circuits could be made to render the output of an amplifier insensitive to varying load impedances; they would deliver the same output voltage level almost regardless of their termination impedance. This eliminated termination problems with the attendant worry of compensating in level for differing load hookups. With the exception of long line feeds, 600 Ω terminations were as good as dead. High-level balanced inputs were now almost exclusively bridging; they had a sufficiently high impedance (usually >10 kΩ) not to disturb the level of the source to which they were tacked on. For better or worse, it has become the conventional studio interconnection technology. It has taken until fairly recently for a distinction and separate level specification for the two technologies to be accepted.

25.6.2 Level Specifications

The original transmission line level specification referred to a power level of 1 mW regardless of impedance. This was 0 dBm. It was a universal specification applicable to any signal of any frequency being transmitted along any length of wire for any purpose at any rated impedance, and it is used extensively in radio-frequency work and other things entirely unrelated to audio. The dBm definition is sacred and can't be changed. Zero dBm in a 600 Ω load works out to 0.775 Vrms; this was adopted de facto as the reference for use in general audio work. With zero impedance technology, although the working voltage is specified, the impedance isn't. It can be anything, but the power (as measured in dBm) necessarily varies as a result; for instance, 0.775 Vrms across a 100 Ω load is +7.78 dBm while across 10 kΩ it would be minus 12.22 dBm. But it's still 0.775 Vrms.

The reference level for zero impedance thinking is a voltage, and the one chosen is the familiar 0.775 Vrms with which everyone was historically used to dealing. That voltage is distinguished as 0 dBu. Some have tried to impose a universal reference based around a voltage level of 1 V called the dBV for audio, which is easily divided by 10 but has proved sufficiently confusing to anyone brought up on the dBm that it is now all but dead.

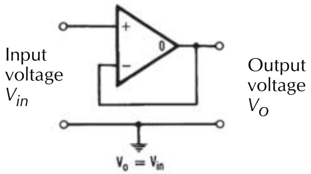
But wait! There's more! The ubiquitous VU meter when implemented as intended imposes a nominal system level of +4 dBm (0 VU = +4 dBm at 600 Ω), and in territories and market segments where the VU reigned, +4 dBm (and latterly +4 dBu) is still a common reference. And try as one might, it is impossible to ignore that there is more semipro recording and audio gear in use than real audio equipment and that generally uses the domestic level of −10 dBu as a nominal reference. Glad that's all cleared up, then.

25.7 Operational Amplifiers in Consoles

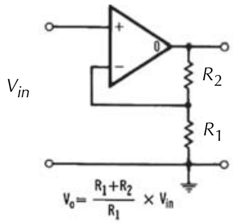
Consoles utilizing integrated circuit operational amplifiers (IC op-amps) have suffered from a curious syndrome, collecting in earlier days a (sometimes deserved) dreadful reputation, which has stuck. This section is an attempt to explain the history, shortcomings, and attributes of IC op-amps from conception to present day, to point out how some shortcomings are overcome and to provide reassurance that they are the future of consoles. It is also an example that this, along with most other technology, is well understood and quantified, the concepts if not the details having been defined many years ago, Fig. 25-16.

When ICs first came out (the Fairchild μA709, e.g., they were expensive, prone to oscillate, and had no short-circuit output protection.

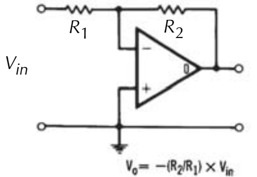
At this stage in the game, discrete transistor circuitry ruled supreme in pro-audio while considerable vacuum-tube gear was still in use. Techniques expanded and ICs were tamed sufficiently to remain operationally stable, but little high-frequency loop gain remained to guarantee enough feedback to adequately reduce high-frequency distortion. Also, they were very noisy. Although their parameters could be set up to be acceptable for any set application and gain setting, the very nature of control in consoles is variable, so the devices almost inevitably ended up operating away from their optimum.



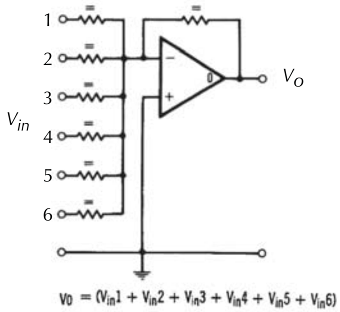
A. Voltage follower.



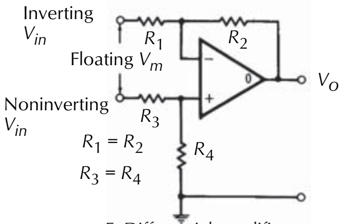
B. Noninverting amplifier.



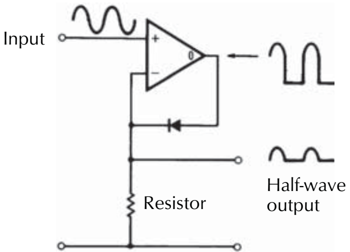
C. Inverting amplifier.


=
1
2
3
4
5
6
V_{in}
−
+
0
V_O
$V_O = (V_{in}1 + V_{in}2 + V_{in}3 + V_{in}4 + V_{in}5 + V_{in}6)$

D. Virtual-earth mixer amplifier.



E. Differential amplifier.



F. Basic half-wave precision rectifier.

Figure 25-16. Basic op-amp configurations.

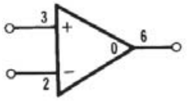
A new term entered the audio design vocabulary: *compensation*. Compensation is the brutal slowing down of the amplifier in order to stop rampant, screaming instability. Essentially it was accomplished by defining the bandwidth of the overall loop around the amplifier or a particular gain stage within the amplifier—or both. And typically robbing the device of its promise.

Hot on the heels of the µA709 came the now much loved and despised, but always revered µA741. Best known in its plastic encapsulated eight-pin dual-in-line configuration, it still took our industry many years to catch on to the fact that here existed a seemingly almost vice-free op-amp. Well, at least, it was free of some of the 709's vices. It was heavily internally compensated to nominally guarantee stability, but the penalty for this was rapidly disappearing open loop gain with increasing frequency. There was just enough gain left to squeeze 20 dB of broadband gain safely over a 20 kHz bandwidth. Some IC manufacturers came up with good 741s, usably quiet and free of the grosser output offset voltage problems that plagued earlier devices. The 741 was also output-protected to the extent of being short-circuit proof, a relief to all.

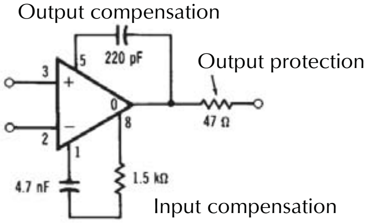
Subsequent generations of op-amps to the 709 included the 748 (the uncompensated sister to the 741) and the 301, again, some versions being excellent for this class of device. That the 748 and 301 were user compensated did allow for more optimal parameter setting and in most circuits only required one capacitor to achieve this (as opposed to the necessary two resistor-capacitor networks for the 709).

Although on the surface this appeared to be of great convenience to the designer, it disguised the fact that far superior bandwidth and phase-margin performance could be obtained by carefully considering the nature of the compensation network. Rather than just a simple capacitor of sufficient value to hold the amplifier stable (which also turned the internal compensated transistor into a Miller integrator doing absolutely nothing for the speed of the device), a more complex network such as a two-pole resistance-capacitance network, Fig. 25-17E, improved matters greatly.

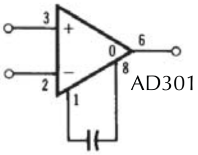
External feed forward, while in use as an inverting or virtual-earth mixing stage, also enabled a dramatic increase in bandwidth and speed over the more conventional compensation arrangements, as shown in Fig. 25-17.



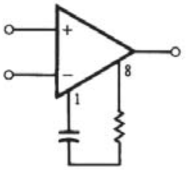
A. Naked op-amp.



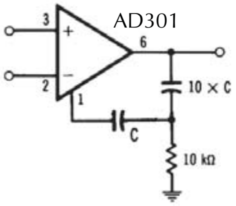
B. Typical 709 compensation.



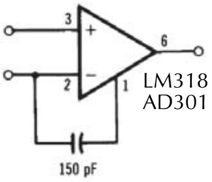
C. Simple single-C compensation reduces slew rate.



D. Single-C with limiting resistor extends bandwidth.



E. Two-pole compensation improves stability and bandwidth.



F. Feed forward considerably improves speed.

Figure 25-17. Various op-amp compensation techniques.

25.7.1 Slew-Rate Limitations

All these early devices had one great failing that has been leaped on vigorously by the hi-fi fraternity and audio engineers alike. Slew rate is the speed (measured usually in volts per microsecond, V/µs) at which an amplifier output shifts when a step source of extremely high speed is applied to the input. All the early-generation op-amps had slew rates on the order of 0.5 V/µs, but no one really understood it or its implications or effects then, and it was not the issue it is now.

Why is slew rate a problem? If the audio signal that the device is attempting to pass has a rise-time that exceeds the amplifier's slew rate, then obviously distortion is created—the amplifier simply cannot react quickly enough to follow the audio. Slew is an issue at high frequencies (rapid transitions) and high levels (a quiet signal is moving at fewer V/µs than the same signal louder). This by happy accident means that a lot of program material that does not contain high amplitudes at high frequencies can be passed by slow, low-slew-rate devices with impunity. Unfortunately, a lot of sources found in recording studios or on stages don't fit that bill; they're loud and have serious high-frequency content.

The speed limitation was nearly always in the differential and dc level-shifting stages of the devices. It is quite difficult to fabricate on an IC wafer ideal classes of transistors in configurations necessary to improve matters without compromising other device characteristics (such as input bias current, which affects both input impedance and offset performance).

Feed forward, in which a proportion of the unslewed input signal is fed around the relatively slow-responding lateral pnp stages, improving slew rate and bandwidth appreciably, is used to great effect in the LM318; a slew rate of some 70 V/µs is achievable by this technique. It was in this area of slew rate, combined with a significantly improved noise performance (again another parameter suffering from difficulty in fabricating appropriate devices in a relatively dirty wafer), that the next major breakthrough occurred in devices commonly used for audio applications—the Harris 911. Although dramatically improved, the slew rate was still not fast and was also asymmetrical (+5 and −2 V/µs).

25.7.2 Bipolar Field Effect Transistors (BiFETs)

A breed of op-amps called BiFETs, or *bipolar field effect transistors,* emerged. These devices have a closely matched and trimmed field effect transistor input differential pair (hence, the typically unimaginably high 10 MΩ input impedance) and a reasonably fast 13 V/µs structure. These devices are typified by the TLO series from Texas Instruments, Inc. and devices such as the LF356 family from National Semiconductor Corp. Selected versions can, when source impedance is optimized, give noise figures better than 4 dB at audio frequencies, which is thoroughly remarkable for units costing very little more than a 741.

The speed of the devices has been achieved by the replacement of the conventional bipolar transistor differential input and level-shifting circuitry with FET configurations. Incidentally, the intrinsic noise characteristic of these FET front ends is significantly different from that of bipolars and seems perceptually less objectionable.

There are currently a few devices designed specifically and optimized totally for inclusion in high-quality audio equipment. With a quoted noise figure of better than 1 dB at audio, a slew rate of 13 V/µs, and the ability to drive a 600 Ω termination at up to +20 dBm, the Signetics Corp. NE5534 (or TDA1034) was in the vanguard of these; many nice devices have since followed.

These are all somewhat more expensive than the BiFET types, but none is prohibitively so, unless the target design is extremely cost sensitive. Today's designer is spoiled by the ability to choose appropriate devices for each application almost regardless of cost, and as will be shown, sometimes the less grand and glorious parts are sometimes the better choice.

Noise in any competently designed and operated console can be attributed mostly to two sources:

1. Mixing amplifiers with an appreciable number of sources and, hence, a lot of makeup gain
2. The input stage, especially a microphone amplifier with a fair amount of gain in it

Once a background noise level is established from the front-end stage (at a level obviously dependent on the amount of gain employed there), the difference in noise contribution further down the line between an amplifier with a typical unity gain noise of −120 dBu and one of −115 dBu is for the vast majority of considerations totally insignificant. Concentration on these two hot spots will define the noise performance of an entire console.

In circumstances where extremely low system noise floors are actually necessary (rather than just deemed a good idea) and where such a noise level isn't being totally swamped by the source (which it usually is), then devices like the 5534 make sense elsewhere. Not so much that they are that much quieter within themselves but that their substantial output driving capability allows circuit impedances to be reduced, resulting in a worthwhile difference to noise floor. It's nice to know that there is also maybe a chance of chucking enough current at capacitors in filters for them to work properly at high frequencies and high levels. Using the 5534 as a microphone amplifier far outweighs the hassle of a similarly performing discrete transistor design, which in this specific area is still its main close rival.

Every design case demands a long cool look to determine what sort of device makes most sense; there is no one-fix cure-all technique, or device. For the most part, the designs here are based on TLO-class or 5534-class devices, determined mostly by whether low-noise, high-output drive capability or high-input impedance driving criterion. Modern devices (for a mature technology, new op-amps do seem to pop up almost weekly) that fulfill these specific or other niche parameters would of course be applicable.

25.7.3 Discrete Operational Amplifiers

The JE990, designed by Deane Jensen of Jensen Transformers and manufactured by Hardy Co. of Evanston, Illinois, is an example of an encapsulated discrete amplifier module. Many fascinating solutions to op-amp internal-design problems (some of which even IC designers evidently haven't realized existed) are implemented in this design whose features demand a total reappraisal of contemporary audio circuit design and philosophy. Optimum input source impedance (normally about 10 kΩ with most IC and discrete amplifiers) is reduced to about 1 kΩ by the use of an IC multiparallel input transistor differential pair. Small inductors in the emitters provide isolation from potential high-frequency instability due to the gain-bandwidth characteristic of the first differential stage shifting with varying source impedances. Unity-gain noise is a quoted staggeringly low −133.7 dBu, while the output is capable of delivering full voltage swing into a 75 Ω load. This permits the use of exterior circuit elements of far lower impedance, reducing thermal noise generation. This elegant device inevitably carries a high price tag. Its many attributes point to the direction for design. It is well ahead of any devices available in IC form and also, to the author's knowledge, of any universal discrete circuitry elements used to date in console manufacture. This device begs the question of the wisdom of the complex multiamplifier, multistage mixer configurations versus true minimum-path circuit philosophy.

25.7.4 Instability

An unexpected thrill facing designers as they upgraded to newer, much faster devices was the tendency for all their previously designed circuits to erupt in masses of low-level instabilities even in what had been perfectly tame boards.

Layout anomalies, such as track proximity, were a major contributor toward the stability problems, so new layouts had to be generated with a whole new set of conditions added to the already hazardous game of analog card design. However, the real roots to this problem are with the devices themselves and a lack of appreciation of the relationship between their internal configurations and the outside world. Everyone who had been brought up designing around 741s had become too used to treating them in a somewhat cavalier fashion and for good reason. It was very hard work to make them misbehave or even show a hint of oscillation. People got used to treating ICs as plug-in blocks of gain with little consideration for the fact that inside was a real, live collection of electronic bits that still had all the problems real electronics always had. The reason the 741 was relatively impervious to user-inflicted problems is analogous to the fact that it's quite difficult to get anything that is bound, gagged, and set in molasses to not behave itself.

Mistake number one with the new devices was believing that they were unity gain stable because the data sheets said so. What that really means is "does not burst into oscillation at unity gain (under these circumstances …)," which is not the same thing at all.

25.7.5 Phase Margin

It is important to maintain as large a margin as possible between the internally structured gain-bandwidth roll-off set for open loop and the roll-off around the external circuitry determining the closed loop gain. This is to preserve sufficient phase margin at all frequencies for which the circuit has gain. Failure to do this can result in the feedback being shifted in phase sufficiently to become reverse phase to that intended (positive feedback) with oscillation resulting. Even if the phase isn't shifted quite that far, the feedback tends toward positive and damped ringing when transients hit the circuit ensues. Also, these resonance effects are extremely high in frequency, typically many megahertz, so any radio signal that gets as far as the circuitry will absolutely adore an amplifier that is critically resonant at its frequency! A reasonable phase margin to aim for at all gain frequencies is better than 45°. In practice, a compromise between desired circuit bandwidth traded off against the need to tighten that bandwidth for the sake of phase margins can be fairly easily reached with the newer devices, provided the need to do so is recognized.

There seem to be two schools of thought on bandwidth versus stability phase margin. First there are the

Pragmatists, who close down the bandwidth of an amplifier as rapidly as possible outside the required passband, maximizing stability phase margin and RF neutrality. Then there are the Purists, who maintain circuit gain as far out and as high as possible, walking the tightrope of stability—usually in deference to the in-band phase linearity.

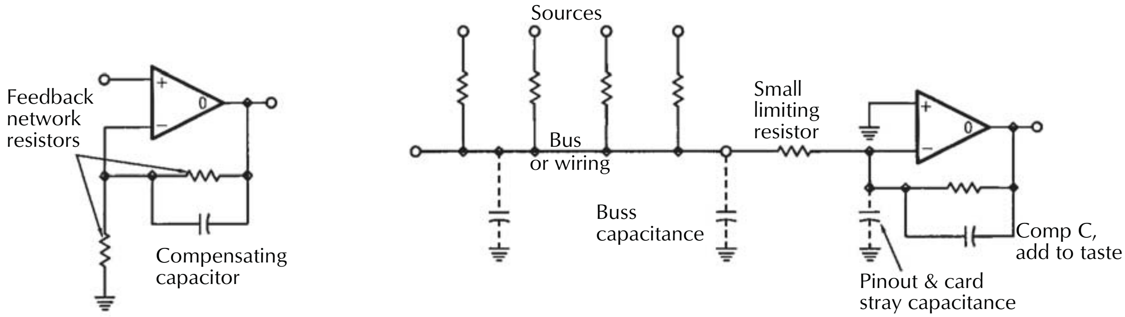
The normal, easiest, and most flexible way to determine the closed loop roll-off of a circuit is by means of a feedback phase-leading capacitor across the main output-to-inverting-input feedback resistor. A typical arrangement is shown in Fig. 25-18. Generally, the need to properly define the bandwidth of a gain block by just such a means automatically takes care of the matter, although it's dangerous design practice to assume that the two requirements—phase-margin determination and bandwidth limitation—are always mutually satisfiable.

A fairly common eroder of phase margin and progenitor of instability is stray capacitance from the inverting input of the amplifier to ground. This capacitance, a combination of internal device, pinout, and printed-circuit layout proximity capacitances, reacts against the feedback impedance to increase the closed loop gain at high frequencies. In normal circuits, even the typical 5 pF or so is enough to tilt up the closed loop gain parameters, threatening stability. Far worse is the situation where the inverting input is extended quite some distance along wiring, and worse yet, a bus—as in a virtual-earth mixing amplifier—hundreds, and sometimes thousands, of picofarads may be lurking out there. It can arise that despite a sizable time constant being present in the feedback leg, none of the expected high-frequency roll-off occurs since it is merely compensating for the gain hike created by bus capacitance. Ensuring required response and phase characteristics using any virtual-earth mixer can only be done properly with at least two orders of compensation around the mix-amp and with the finished system up and running completely, since any additional sources modify the impedance presented by the bus.

To define just how much this unwanted gain can rise, a small limiting resistor may be added as close to the amplifier inverting input terminal as possible; this is at the expense of the virtual-earth point now having a minimum impedance based on the value of that resistor. The resistor, incidentally, is also a measure of protection against any radio-frequency signals on the bus being rectified by the input stage junctions. Better yet, a small (real!) inductance in series with the summing amplifier input provides another means of out-of-band gain reduction and RF immunity.

25.7.6 Time-Domain Effects

There is invariably a finite time taken for a signal presented at the input of any amplifier to show an effect at the output of the amplifier—the so-called transit time. Every tiniest capacitance and consequent time constant in the internal circuitry of the amplifier make this inevitable; electronics takes time to do things. This transit time becomes an appreciably greater proportion of the wavelength of the wanted signal as the frequency increases, and as such it has to be taken into account. Fig. 25-19 shows how the fixed transit time becomes more relevant to increasing signal frequency. Ultimately, of course, the transit time will become half the time necessary for a wavelength of the signal frequency. At that stage what emerges from the amplifier will be a half wavelength or 180° out of phase. Before this point, its detraction from phase margin with increasing frequency can start to cause serious problems; at this

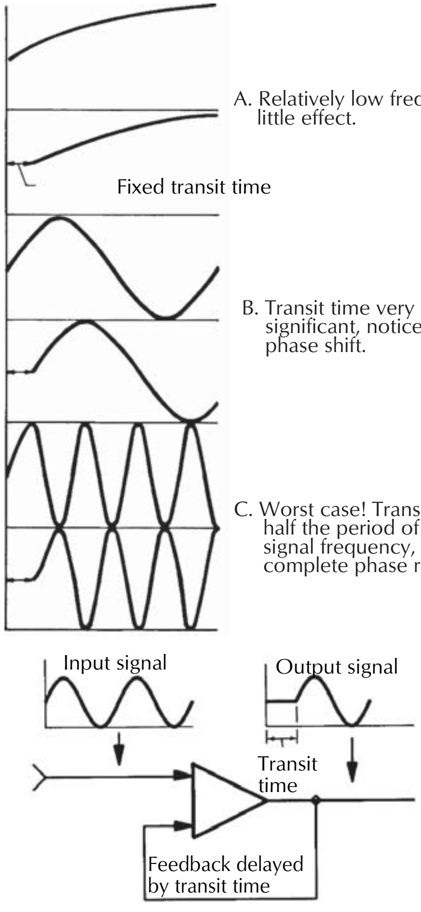


A. Basic external feedback compensation

B. Virtual-earth mixer with stray pinout and bus capacitances.

Figure 25-18. Feedback phase-leading stability compensation.

ultimate state, though, the negative feedback on which the amplifier depends for predictable performance is now completely upside down. Now it's positive feedback. Now the amplifier oscillates.


A. Relatively low frequency, little effect.
Fixed transit time
B. Transit time very significant, noticeable phase shift.
C. Worst case! Transit time half the period of the signal frequency, complete phase reversal.
Input signal
Output signal
Transit time
Feedback delayed by transit time

D. Effect in feedback circuit.

Figure 25-19. Transit time effects with increasing signal frequency.

25.7.7 Transient Intermodulation Distortion (TID)

The TID effect, if not fallout from and overwhelmed by the effects of insufficient slew rate, is due to amplifier transit times. Not surprisingly, as is nearly always the case with fad problems (as was TID during the 1970s), TID has been known and appreciated for as long as there have been negative feedback amplifier circuits—the twenties. It is and always has been totally predictable.

TID is a direct result of the servo nature of an amplifier with a large amount of negative feedback. The feedback is intended to provide a correction signal derived as a difference between the amplifier output and the applied input signal. It is a simple concept: any difference between what goes in and what comes out is error in the amplifier. All we need do is subtract the error. However, it is not so simple. Since there exists a time delay in the amplifier, the circuit has to wait for that amount of time before its correction signal arrives. The output during this time is uncontrolled and just flies off wildly in the general direction the input tells it to. Once the correction arrives, the amplifier has to wait again to find out how accurate that correction was and so forth, see-sawing on and on until the amplifier output settles. Fortunately, this all takes place rapidly (depending on the amplifier external circuitry), but it still represents a discrepancy between input and output. It is an effect peculiar to amplifiers with large amounts of negative feedback (typical of most contemporary circuitry), frequently displaying itself quite audibly—especially in power amplifiers where transit time is quite long with the usual huge, slow output devices.

Amplifiers that rely on their own basic linearity—such as tube amplifiers—rather than on a servo-type nonlinearity correction system, are often held to be subjectively smoother. A whole subindustry thriving on the virtues of feedbackless circuitry has evolved. Nowadays, though, with device speeds improved as they are, settling times are becoming insignificant in relation to the signal transients with which they are expected to cope, pushing the frequency area at which TID could manifest itself far, far beyond expected audio excitation.

25.7.8 Output Impedance

A lot of devices, particularly the TLO series of BiFETs, have a quite significant open-loop output impedance. This is because the IC designers obviously considered that instead of an active output current-limiting circuit (standard on most op-amps up until then), a simple resistor would suffice. Although this built-in output impedance—by virtue of the enormous amount of negative feedback used—is normally reduced to virtually zero at the output terminal, it is still present and included as part of the feedback path, Fig. 25-20. Any reactive load at the output is going to materially affect the feedback phase and phase margin.

Any capacitance from the output to ground will form a feedback phase-lagging network. This shifts the phase

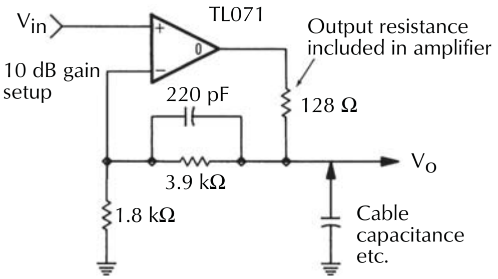


Figure 25-20. Output impedance as part of the feedback loop.

inexorably toward the point where the total amplifier and network phase shift reaches 180° at the inverting input (that's a full 360° total), and the circuit oscillates. The frequency at which it oscillates is inversely relative to the capacitance value. It isn't unusual, with small values, to find oscillations right at the edge of the high-frequency sensitivity of an oscilloscope. Hanging a long piece of wire on the amplifier output (especially shielded cable with its high shield to inner capacitance) is a surefire guarantee of instability for this very reason. It has the added complication that there is a measure of inductance there, too. It is conceivable that a long cable might start to look like a mismatched tuned stub at a frequency where the amplifier still has some gain, creating a creditably good, stable RF generator.

What this extra resistance-capacitance output circuit is in effect doing is to add dramatically to the transit time of the amplifier where actually the termination problem is creating far more delay than could possibly exist within the device itself. That the cures for the two ills are similar shouldn't be a surprise. Fortunately, a simple fix for this instability is to buffer away the load from the output feedback termination with a small resistor of typically 33– 150 Ω. This usually does it, but at the expense of head room loss due to the attenuation from the buffer resistor against the load termination. Provided the load is greater than about 2 kΩ, which it would really have to be in order to prevent getting close to current drive saturation in the IC output stage, this head room loss should be well less than 1 dB. A better way is to buffer off with a small inductance, giving increasing isolation with frequency; a phase-shifting characteristic opposite to that of the (normally) capacitive load provides a total termination that is phase constant at the higher frequencies. At the lower audio frequencies, of course, the inductive reactance is very low, and the load sees the very low dynamic output impedance of the amplifier. The buffering inductance becomes virtually transparent.

Both of these techniques also provide a measure of protection against the possibility of RF signals finding their way into the amplifier by means of rectification in the output stage or inverting input. Very often output stages are more prone to RF field detection than inputs.

Some devices with a quite low output impedance before applied feedback (i.e. those with unbuffered, complementary emitter-follower output stages) are not likely to be fazed as much by these effects (pun totally intentional) but it is just as well to design in these considerations habitually. Emergency replacement, device upgrades or IC internal design changes can evoke this problem unintentionally.

25.7.9 Compensating Op-Amps

Op-amps generally have a couple of pins dedicated for compensation, which can be taken as a less than subtle message from the manufacturer that their product isn't stable under certain conditions of usage and needs external kludging. Usually this is at low closed loop gain where the bandwidth is at its most extreme. The classic solution is to shrink the bandwidth of the amplifier by slowing the amplifier down. Among other things, this wrecks the slew rate that's been handsomely paid for.

The most ordinary means of slowing down the devices is to slug an internal gain stage, leaving the other stages intact. On the bright side, if it is this internal gain stage around which the external compensation capacitor is hung that is tending toward instability, the capacitor should cure it. Sadly, it rarely is that stage. If a previous stage, say, the input differential amplifier, is unstable, all the capacitor will do is slow up the amplifier and reduce the slew rate to the extent that the oscillation is no longer visible at the output. It does not cure the instability. It's still in there, hiding. Often the only external manifestations are supposed dc offset voltages that won't go away and a poor-sounding amplifier.

There is a moral to this tale of compensation: don't use op-amps that require compensation if at all avoidable. Stability should be ensured by the circuit as a whole, and if speed is to be preserved, the op-amp should not be used below the gain at which it's happily stable. *Compensation achieves stability by masking a symptom and not by tackling the cause.*

The previous precautions, in addition to the feedback phase-leading capacitor, are now required circuit practice for using the newer, fast devices in many op-amp configurations. It should be said here that because there is no facility for implementing phase leading around the standard voltage-follower configuration and that this is the most critical configuration for stability, it is not a preferred circuit element. The manufacturer will have designed the IC to be just stable enough at unity gain to be able to say so unblushingly, but with probably little real-world margin to spare. Hanging a compensation capacitor across the appropriate pins will slow up the slew rate and not necessarily make the whole amplifier any less unstable. It is better not to tempt fate.

25.7.10 Input Saturation

The use of a standard voltage follower implies that in order to maintain the same system head room in that stage, the input has to rise and fall to the same potentials that the output is expected to. It can't. In most op-amps, especially those with bipolar inputs, the differential input stages saturate or bottom significantly before the power supply rails are reached and certainly before the output swing capability is attained. This limited input common-mode range means that the follower not only will cease to follow but will also spend a considerable amount of time in unlatching from one swing extreme or the other. Once an amplifier internal stage has latched, the feedback loop is broken; the stage has no assistance from the servomechanism to unstick itself. Once the loop is reestablished, it has to settle again as if from a hefty transient before it can resume following. Basically, this is an ugly scene. Uglier yet is the propensity of some devices when the input common-mode range has bottomed for the output to lunge to the opposite rail. Talk about "sonic character."

IC manufacturers commonly specify the common-mode input voltage range and it is precisely this limit that would be exceeded in use as a follower. For reference they are: ±13 V for the 5534, ±11.5 V for an LM318, and +15 V to −12 V for a typical BiFET.

All fall far short of the power supply maxima. Provided enough gain is built around the amplifier to prevent these common-mode limits from being reached, there should be no latching hangups; the feedback network also provides some substance to hang closed-loop compensation around in addition to enabling the full output voltage swing of the amplifier to be utilized.

Similar settling-time problems occur any time any stage is driven into clipping, but given the high power-supply voltages and consequent large head-room common today, clipping should be rare.

In short, not only for this good reason, the standard voltage-follower configuration is pretty bad news.

25.7.11 Front-End Instability

Altogether the most obscure potential instability-causing effect relates directly to the behavior of the input stage in bipolar front-end op-amps. The gain-bandwidth characteristic of the input differential stage is greatly dependent on the impedance presented to the input, the gain-bandwidth increasing with reducing source impedance. There is the possibility that given an already critical circumstance, the erosion in phase margin due to this effect can cause overall instability. This instability can be mitigated by limiting the gain-bandwidth excursion by means of a resistor (typically 1 kΩ) and/or some inductance in series with the input. Ordinarily, this would have little effect on circuit performance but may, especially in microphone amplifiers, detract from noise performance. Noise performance is largely dependent on the amplifier being fed from a specific source impedance, and 1 kΩ would be a sizable proportion. However, it's usually fairly easy to arrange in the design stage such that the IC doesn't have a zero impedance at either of its inputs.

Fortunately, because of the far greater isolation between the FET gates and their channels, this is a problem that FET-input op-amps do not have. A similar approach to that proposed for output isolation (i.e., an inductor rather than a resistor) in series with the affected input seems, on the surface, an equally good idea. The impedance of the inductors would be low at audio frequencies (so not affecting noise criteria significantly) and high at radio frequencies where the low source impedance phenomenon does its work. Unless the value is critically defined, an inductor of sufficient value to provide a usefully high reactance at RF also could be self-resonant with circuit stray and its own winding capacitances at a frequency probably still within the gain-bandwidth capability of the amplifier. Takes a bit of care.

Those who have experienced design with discrete circuitry will not be surprised that this source impedance instability effect is also the reason emitter followers are the most instability prone of the three basic transistor amplifier configurations. The cure is the same. Not only does the series impedance limit the source impedance before zero, it also acts together with any pinout and base-emitter capacitance as a low-pass filter helping to negate further external phase shift that may detract from stability. This base source-impedance instability is quite insidious in that it can either contribute to instability of the amplifier loop if it is already critical or it can be a totally independent instability local to the affected devices with nothing whatsoever to do with the characteristics of the external loop.

25.7.12 Band Limiting

One of the first great superficially appealing results of using the enormous feedback inherent to op-amps at the relatively low gain requirements of the audio world was a close approach to dc-to-light frequency response. The author remembers well the hysterical peals of laughter as the response of a new mixer was measured as still 0 dB right to the end of the testing ranges of the oscillator and the badly disguised puzzled looks and worried glances when we put real audio through and actually listened to it.

Many audio signals, especially live ones from microphones, analog tape-machine returns with a high vestigial bias content, keyboards, and a range of other sources, have a fair amount of ultrasonics present. If an analog remix of a digital recording takes place (it happens) many digital-to-analog (D/A) converters have an embarrassment of out-of-band noise that is of no program relevance whatsoever. A good microphone is going to hear all manner of stuff in a space: TV/computer scan whistles, motion alarms, switch-mode power supply or light-dimmer inductor screechings just for a start, none of which can be pretended to be musical. Depending on how good a following A/D convertor implementation may be, some of these may well get aliased down into the audible frequencies to less-than-subtle effect.

There is a proverb: the wider the window, the more muck flies in. Returning focus here just to analog signal processing, it would be perfectly all right if the following circuitry were capable of dealing with signals much higher than the audio band; sadly at the time (and to a lesser degree even now) that is not so. The root of the difficulty is the worsening open loop gain of the individual op-amps; as it drops off at 6 dB/octave with increasing frequency, there remains less closed loop feedback available to maintain the op-amp's linearity. In other words, the circuitry becomes less and less linear as the frequency increases and the feedback dwindles.

Fig. 25-21 is representative of the open loop (no feedback) input-output transfer characteristic of an op-amp—i.e., what comes out in relation to what goes in. Not at all linear. In fact, rather nasty. (Incidentally, most big power amps have similar curves.) The good in-band linearity and low distortion of op-amps come from the application of monstrous amounts of negative feedback. Take the case of a noninverting 741-type amp with 40 dB of gain around it, Fig. 25-22. At 100 Hz there can be 60 dB of feedback, which is great nonlinearities are being corrected by roughly the tune of 1000:1! However, the open loop gain plummets above this frequency, leaving a still respectable 40 dB of feedback at 1 kHz (100:1). (This figure of 40 dB is widely regarded as the lowest amount of feedback for good performance from an op-amp.) At 10 kHz it's down to 20 dB; it is 14 dB at 20 kHz; and at an ultrasonic 40 kHz, there is a bare 8 dB! There is still gain, though, and the amplifier is quite capable of supporting and amplifying a signal up at those frequencies; it's just not very good at it.

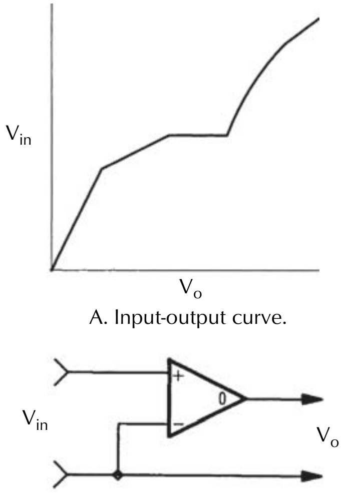


Figure 25-21. Operational amplifier open-loop gain curve typical of a bipolar device.

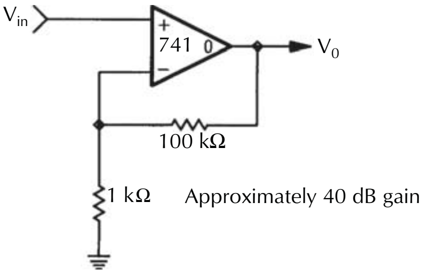


Figure 25-22. A 741 with 40 dB gain.

Harmonic distortion of ultrasonics that would be generated by passing through a transfer function like Fig. 25-21 is unimportant; the frequencies would be even more ultrasonic. The problem lies in the intermodulation of two or more signals, products of which more often than not fall into the audible band; even reciprocal mixing with noise results in in-band noise products. A whole slew of intermodulation products are produced. It is no wonder that early op-amps sounded bad.

So much for the expected result of improved transient response through having a wide-open frequency response. As is now obvious with hindsight, deliberately limiting the input frequency response of the mixer to a little more than the audio band results in an amazing cleanup of the sound. By removing a lot of the inaudible signals that cross modulate within themselves and with in-band signals, the cause of much of the lack of transparency and mush that had become the trademark of early-generation IC op-amp consoles is eliminated.

Despite improved devices with greater open loop gains at far greater bandwidths, this approach remains valid today. By band limiting the program signal to reduce inaudible signals as early in the chain as possible, there is far less chance of their generating unwanted audible products. A front-end low-pass filter, operating in conjunction with all the other low-pass effects of feedback compensation arrangements throughout the console, should provide adequate minimization of these products in modern devices.

Purist arguments about the undesirability of any deliberate filtering seem rather futile in a world of real devices—all but a very few transducers and what they hear/reproduce are an embarrassment above 20 kHz, and final signal destinations—like anything digital, or otherwise inherently band limited. However, with 96 kHz digital sampling threatening to become mainstream, widening the window to at least utilize some of the fabulously hard-won bandwidth may be in order in systems where such is likely. Band limiting, to whatever sane degree, is a particularly powerful tool for obviating funny noises and lack of sonic transparency, and its use shouldn't be abdicated without a fight.

25.7.13 Slew-Rate Effects

Slew-rate limiting occurs when the fastest signal rise time the amplifier is expected to pass exceeds the speed of the fastest stage in the amplifier; the input transient becomes slurred to as fast (or slow) as the amplifier's capability. It is a level-dependent effect; at low levels the input signal's transient may be well within the amplifier slew envelope and escape unmutilated, but as the input gets larger the transient's slope can equal or exceed that of the amplifier.

Slewing gives rise to intermodulation effects that are dependent upon both frequency and signal level. The louder and faster the input transient, the worse the damage. A common subjective result of this limiting is for the high end of a drum kit to change in character of sound with differing levels of the lower-frequency instruments on which it is riding. Another favorite is the "disappearing snare drum" in which, again, the sound radically alters with changing level.

25.7.14 Device Idiosyncrasies and the Future

Many circuits rely somewhat on the extremely high input impedances of the BiFET devices and their very low required input bias currents. Using bipolars everywhere may result in unavoidably generated output offset voltages that could manifest themselves in extreme instances as switch clunks and scratchy pots. Also, the feedback phase-leading compensation may or may not be appropriate for devices other than BiFETs, especially with some bipolars with less than ideal internal poles. If there's a temptation to use more conventional bipolar devices, particularly those in multiple packages, it is also worthwhile examining their characteristics when inputs or outputs are taken above or below the power supply potentials. If the device structure under such circumstances is unprotected and turns into a silicon-controlled rectifier that deftly shorts the power supply with a bang, you are possibly better off using something else. In short, if a device is chosen specifically for an application and support circuitry designed for it, adding another device for the sake of adding it is usually nonproductive and often a step back. Op-amps and their surrounding components should be regarded holistically.

The proliferation of amplifier elements in modern console design has mushroomed further in recent years with the availability of compact and extremely low-cost IC op-amps. Increasingly complex functional blocks are becoming increasingly commonplace. If, in order to improve their electrical and sonic characteristics, it would mean an increase in size and cost of well over an order of magnitude, would they still be quite as popular? In the good old days of tubes, it was not through any lack of expertise that equalizers even of today's complexity did not exist; it was just the size and cost would have made even the reckless shudder. Also, it is to be noted, they were not really thought necessary.

By way of history repeating itself, though, the astounding complexity of many digital audio algorithms (e.g., the use of as many as nine biquads to achieve not a whole EQ, just a single section, which alone would require a mere 27 op-amps to emulate) makes concerns about analog technology overkill seem a touch quaint.

25.8 Grounding

A human working visualization of anything electronic soon becomes impossible without a mental image of the solid, infinite, immovable, dependable ground. It has many other names too: earth, 0 V, reference, chassis, frame, deck, and so on, each of differing interpretation but all, ultimately, alluding to the great immovable reference.

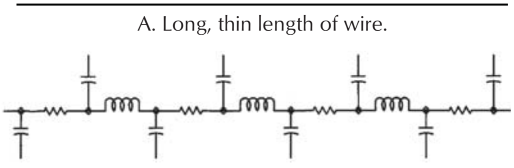
Electrons could not care less about all this. They just go charging about as potentials dictate; any circuit will work perfectly well referred to nothing but itself. (Satellites, cars, and flashlights work, don't they?) Ground in these instances is but an intellectual convenience.

Interconnection of a number of circuit elements to form a system necessarily means a reference to be used between them. To a large degree, it's possible to obviate a reference even then by the use of differential or balanced interfacing, unless, of course, power supplies are shared.

So, having proved that ground is seemingly only a mental crutch, why is it the most crucial aspect of system design and implementation?

25.8.1 Wire

Fig. 25-23A shows a typical, ordinary, long, thin length of metal known more commonly as wire and occasionally as printed-circuit track. However short it is, it will have resistance, which means that a voltage will develop across it as soon as any current goes down it. Similarly, it has inductance and a magnetic field will develop around it. If it is in proximity to anything, it will also have capacitance to it.


A. Long, thin length of wire.

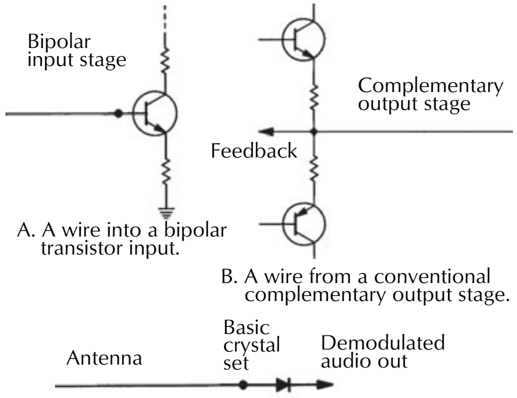
B. The wire actually has resistive and distributed reactive components.

Figure 25-23. What is a length of wire?

So Fig. 25-23A actually looks more like Fig. 25-23B with resistive and distributed reactive components. Admittedly, these values are small and seem of little significance at audio frequencies, but clues have already been laid (particularly in Section 25.7 on op-amps) that believing the world ends at 20 kHz is not so much myopic as naive.

A radio engineer looking at Fig. 25-23B would mumble things like "transmission line," "resonance," or "bandpass filter," maybe even "antenna." RF technology and thinking may seem abstruse and irrelevant to audio design until it is considered that active devices commonly used nowadays have bandwidths often dozens, sometimes hundreds of megahertz wide. An even more frightening realization is the enormous quantity of RF energy present in the air as a consequence of our technological being; never mind the gigawatts of broadcasting bombarding us, the proliferation of walkie-talkies, cellular phones, and business radio all beg mischief in our systems. It even comes from other continents; the aggregate field strength of international broadcasters clutched loosely around 6 MHz, 7 MHz, 10 MHz, and 15 MHz is truly phenomenal.

A more obscure collection of equivalents is shown in Fig. 25-24. Fig. 25-24A represents a wire into a bipolar transistor input; Fig. 25-24B shows a wire from a conventional complementary output stage; and, for reference sake, Fig. 25-24C shows a basic crystal-set radio receiver. It may seem quaint, but for the presence today of wildly more volts per meter RF field energy compared to the heyday of wireless it works just the same. In all the three circumstances, radio frequencies collected and delivered by the antenna are rectified (hence, demodulated and rendered audible) by a diode (the base-emitter junctions in Figs. 25-24A and B). As contrary as it may seem for demodulation to occur at an amplifier output, it is perhaps the most common detection mechanism with the demodulated product finding its way back to the amplifier input by means of the conveniently provided bypassed negative feedback leg.



A. A wire into a bipolar transistor input.

B. A wire from a conventional complementary output stage.

C. A basic crystral-set receiver.

Figure 25-24. A collection of equivalents.

Making our length of wire fatter and thicker has the effect of lowering the resistance and inductance while increasing capacitance (greater surface area exposed to things nearby). So, although the resonant frequency of the wire stays about the same, the dynamic impedance (hence, Q) reduces. Although in general this is deemed a good thing, in some instances it can merely serve to improve the matching and coupling of the RF source to the resonance.

Carried to an extreme, even a console frame constitutes a big fat resonant tank at a surprisingly low (mid-VHF) frequency while frame resistance, however heavily it may be constructed, cannot be disregarded and cannot be treated as a universal ground path.

For the purposes of practical design, these considerations perhaps become a little better defined. The reactive elements of capacitance and inductance with the attendant effects of resonance and filtering are concerned with less obvious aspects (such as electronic stability and proneness to radio demodulation), while resistance gives rise to most of the horrors usually lumped under the collective term *grounding problems*.

25.8.2 Earth Ground

The closest most of us get to earth is the fat pin on an ac power plug. Fortunately for most purposes, it is adequate, provided just the one point is used as the reference. Other points are likely to have slightly differing potentials due to dissimilar routing and resistances. Compared to a technical earth ground (e.g., a copper water pipe or, alternatively, a fortune in copper pipe hammered into the earth), conventional earth grounds can have a surprisingly high potential, a volt or two, considering it is principally a safety facility not ordinarily carrying current. Any potential implies resistance in the earth path, which is bad news about something intended as a reference while also detracting from the safety aspect.

Practically, though, it does not matter too much if everything is waving up and down a bit provided everything, including even unrelated things in proximity, are waving up and down in the same manner. The potential is usually small, meaning that the ground impedance is reasonably low to the extent it may be considered insignificant.

25.8.3 Why Ground Anything to Earth?

With all our component system parts tied together by a reference ground and everything working as expected, the question arises as to why it is necessary to refer our ground to earth. If the internal grounding is completely correct, our system will operate perfectly, quietly, and tamely regardless of to what potential (with respect to earth) it is tied. If not tied, it will derive its own potential by virtue of resistive leakages, inductive coupling, and capacitance to things in its environment. For an independently powered system (i.e., batteries), these leakages and couplings will be of very high impedance and, hence, easily swamped by human body impedance to earth.

If, as is most often the case, most of the system is powered off the ac lines, this floating ground potential becomes of far lower impedance and consequently is much more capable of dragging current through a human load. That's you or me. (It's the current that kills, not the voltage.) A telltale sign is a burring, tingling feeling as you drag a finger across exposed metalwork on something that is deriving its own ground potential.

The mechanism for this lower impedance is fairly straightforward. Power transformers are wound with the optimum transfer of energy at 50–60 Hz and very high flashover voltages (several kilovolts) in mind; the finer points of transformers such as leakage inductance, interwinding, and winding imbalance capacitance are all but disregarded.

Being far greater in scale than ordinary ambient reactive couplings, they primarily dictate the floating ground potential to be anything up to 240 Vac above ground or whatever the power lines happen to be locally.

It used to be that some units were fitted with bypass capacitors from each supply leg to chassis ground, partly in the fond hope that this would help prevent any nasty noises on the ac mains from entering the hallowed sanctum of audio within. Ungrounded, this guaranteed the chassis floating at half the supply rail from a fairly low impedance. Ouch. But it gets worse. With the near universality of switch-mode power supplies, and with nearly everything containing digital electronics to some degree or other, there is the imperative of ensuring any nasties that the supply/digits generates don't find their way out of the box and up the power lead! Some reversal of fortunes there. This is required not so much from the altruistic desire to not pollute but that the box likely wouldn't pass emissions testing (FCC Part 15, CE and the like) and wouldn't be able to be sold. Typical supply-side filtering as a minimum in such supplies is a π filter of common-mode inductors and parallel capacitors—including capacitors to chassis ground.

The result is that if the chassis is not directly earthed, it rides at (in the case of both lines having tied capacitors) half the line voltage. The capacitor values grossly swamp transformer and ambient leakages and give the chassis floating potential an uncomfortably (literally) low impedance. The chassis tingle changes from "Mmm—interesting" to vile oaths with attendant flailing limbs.

A system composed of many separately powered units will almost certainly hum, buzz, and sound generally uneasy if not earthed, which is seemingly in direct contradiction to the earlier statement that "the system will operate perfectly regardless of what potential it is tied to." Being tied to a lot of different self-generated potentials at a lot of different points along a system path is definitely not in the recipe.

Each different power transformer will have different amounts and permutations of leakage and, hence, propagate different potentials and degrees of power-line-borne noise into our otherwise perfect grounding path. Assorted ground potentials mean assorted ground currents, meaning assorted noises.

Tying the entire grounding path to earth is the best shot at "swampout" of leakage impedances. A connection to a (nearly) zero impedance makes nonsense of most other potential-creating paths, most of which have reactances in the kilohms.

Regardless of earth termination in such a multisupply circumstance, significant currents exist along the ground reference lines. The resultant interelement noise and hum voltages (developed across the inevitable line resistances) quickly become intolerable in unbalanced systems. Any wobbling of the ground reference becomes directly imposed upon the desired signal.

Balanced, or pure differential, transmission helps to obviate these perturbances by rendering them common mode in a system that is (theoretically) only sensitive to differential information. In reality, practical transformers can afford a good 70–80 dB common-mode isolation at low audio frequencies. They deteriorate in this respect at 6 dB/octave with increasing frequency up to the winding resonance frequencies unless considerable effort is made to fake a more accurate balance externally. Although transformer balancing does effect a dramatic improvement in noise levels, it is far greater for fundamental hum (50–60 Hz) than it is for other power-line-borne noise. This explains why in tricky systems, lighting dimmer buzz, motor spike noise, or any source with a high-frequency energy or transient content is so persistent.

The golden rule is to treat the grounding of any balanced system as if it were unbalanced. This minimizes the inevitable reference ground currents.

There is one overarching good reason only glanced by earlier for grounding to earth. The consequences of a piece of the gears' chassis becoming inadvertently at the power-line potential are obvious. We would much rather see death to a fuse or breaker than to one of us.

25.8.4 Console Internal Grounding

Let us assume that the grounding for the studio control room is all sensible and that our console has a solid earth termination. What about the intraconsole grounding paths? For most console builders this is perhaps the ultimate unbalanced signal path.

Conventional amplifier stages rely on a voltage difference between their input and reference in order to produce a corresponding output voltage (referred, naturally, to the reference of the input). If the input is held steady while the reference is wobbled, a corresponding (amplified) inverted wobble will appear at the output.

It is plain that any signal the reference sees that is not also common to the input (e.g., ground noise) will get amplified and summed into the output just as effectively as if it were applied to the proper input. The obvious (and startlingly often overlooked) regimen to render extraneous noise unimportant is to ensure that the point at which an amplifier source is referred is tied directly to the reference, while that amplifier output is only taken in conjunction with the reference. Successive stages daisy chain similarly—source reference to destination reference, and so on. This philosophy is called *ground follows signal*.

25.8.4.1 Ground Follows Signal

"Ground follows signal" is a classic maxim and one that has dictated the system design of nearly every console built. It was particularly true in the era of discrete semiconductor design, where ground was often not only audio ground but also the 0 V power-supply return; ideally the audio and supply grounds should be separate. As an added complication, the power-supply positive lines, being heavily regulated and coupled to ground, were an equal nightmare as they too became part of the grounding path. This could be fairly simply avoided by spacing each circuit element away from the supply line by an impedance considerably greater than that offered by the proper ground path—achieved by either separately regulating or simply decoupling by a series resistor, parallel capacitor network, as shown in Fig. 25-25. This actually gives the lie to the notion that single-rail supply systems are easier than differential rail arrangements; to do them properly results in almost indistinguishable numbers of parts and degrees of effort.

Accelerating technology has for once actually made life a bit simpler—specifically, the trend toward IC op-amps with their required differential ($+V_e$ and $-V_e$) power supply. This, thankfully, removes electronic operating current from the audio system ground, while individual stage supply decoupling is rendered a nicety rather than a necessity by the excellent power supply noise rejection ratio of most popular op-amps.

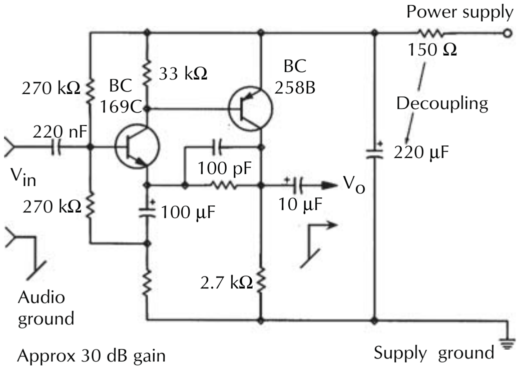


Figure 25-25. Power-supply decoupling: A typical discrete amplifier with the power supply isolated from audio ground.

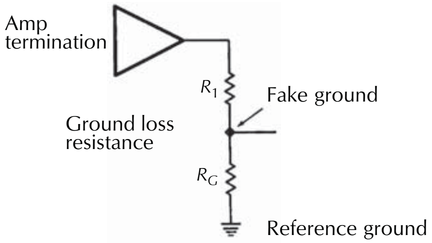
Nevertheless, correct grounding paths still apply; the removal of supply current just exposes and highlights audio ground subtleties.

Unfortunately, although op-amps have simplified matters in one respect, their ease of use and versatility have been largely responsible for the creation of enormous systems with so many stages, break points, mix buses, and distribution networks that the simple daisy chaining of ground follows signal becomes unwieldy, if not unworkable. Alternate grounding schemes, such as star grounding where every ground path and reference is taken to a central ground or earth, play increasingly important roles.

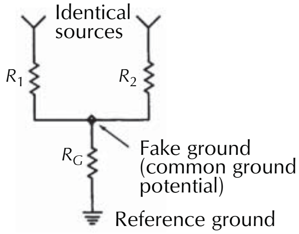
In practice, a necessary compromise between these two prime systems occurs in most console thinking. Daisy chain applies mostly to on card electronics (e.g., in the microphone amplifier sections), while systems switching and routing rely on star connections.

25.8.4.2 Ground Current Summing

A principal grounding-related manifestation is crosstalk, or the appearance in a signal path of things that belong elsewhere. Other than airborne proximity-related reactive crosstalk, the most unwanted visitations are by the common-impedance or resistive ground path mechanism. In Fig. 25-26A, R_1 represents the load of an amplifier output (whether it's the 10 kΩ of a fader or a 600 Ω line termination is immaterial for the present). The resistor R_G represents a small amount of ground path wiring, loss resistance, and so on. It is quite apparent that the bottom end of the termination is spaced a little way from reference ground by the wiring resistance, and the combination forms a classic potentiometer network. The fake ground has a signal voltage present of the amplifier output voltage attenuated by R_1 into R_G.



A. Load of an amplifier output.



B. Two terminals sharing the same ground.

Figure 25-26. Ground current summing.

Practically, with a 600 Ω termination (R_1) and a ground loss (R_G) of 0.6 Ω, the fake ground will have a signal voltage some 60 B down. The use of the fake ground as a reference for any other circuitry is a surefire guarantee of injecting –60 dB worth of crosstalk into it.

Two identical terminations sharing the same fake ground, Fig. 25-26B, happily inject a small proportion into each other by generating a common potential across the ground loss R_G.

Should the second termination be far higher in impedance (the 10 kΩ of a fader), its contribution to the common fake ground potential will be far less (–86 dB) since the ground impedance is much smaller in relation to the source. Correspondingly, though, this higher impedance termination is more prone to be crosstalked into from the lower impedance contributors to the common ground.

25.8.4.3 Typical Grounding Problems

Here is a fairly unusual (but definitely not unknown) grounding anomaly resulting from inattention to the grounding paths. In Fig. 25-27 A2 is a line amp feeding a termination of 600 Ω into a lossy ground of 0.6 Ω resulting in a fake ground potential 60 dB below the output of the amp. An earlier stage in the chain A1 (in this example, a microphone amplifier, with a considerable amount of gain) has its feedback leg (amplifier reference) tied to the same fake ground. Its input ground reference (here lies the problem) is taken from a separate bus supposedly to provide a nice, clean ground. This it does admirably, the bus being tied straight to reference ground and having no sources of great substance going to it.

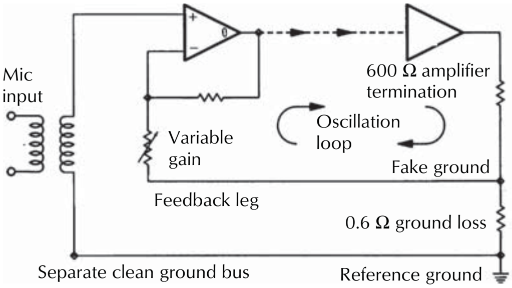


Figure 25-27. Feedback and oscillation via poor grounding.

Any signal present on the fake ground is duly amplified by the microphone amplifier (in its inverting mode) and is attenuated at the line amplifier output back into the fake ground. Naturally, as soon as the microphone amplifier gain exceeds the output attenuation, the entire chain bursts into oscillation.

A very similar mechanism was responsible for an owner's criticism of his well-known console that whenever he attempted to use the track routing on any channel modules, the sound of that channel discernibly altered. It was found that ordinarily nothing in the channel drew much current; all ground impedance requirements were quite light. Light, until the track routing line amp with its load of routing resistors and a terminated output transformer was accessed, demanding a relatively large ground current. This output stage current shared the only ground access point of the module (two paralleled connector pins) with all the rest of the module electronics, with the notable exception of the microphone and line input transformer ground returns. The resultant feedback, although nowhere near enough to promote oscillation, did by virtue of the phase shifting of the output transformer at both high and low frequencies result in distinct coloration.

A purist answer to these fake and loop problems is to choose one grounding point for the entire console and to take every reference and ground return directly to it through separate ground wires.

A few less than minor problems would ensue. The enormous number of ground lines would soon outstrip the capacity of the module connectors, and the mass of wiring would cause apoplexy from the wiremen and aggravate an already critical world shortage of copper. Fortunately, a working compromise suggests itself based on separating the different classes of ground requirements by impedance.

Bucket grounding refers to tying fairly high-impedance sources to a common ground point, bus, or line (since the ratio of their impedances is so great that resultant fake ground potentials can hopefully be made low enough to ignore). Anything that is likely to draw current (any kind of output or line amplifier stage) should go directly to ground, will not pass through any bus, and will not collect shared ground paths on the way to the bucket.

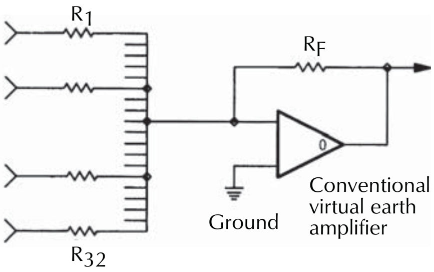
Any ground bus will have a measure of resistance and must, therefore, be fake to a certain degree. If we do our sums right, ground bus signal levels can be kept acceptably low, below −100 dBu.

Smugly, we can expect to ignore figures like that until we (almost inevitably) amplify them up.

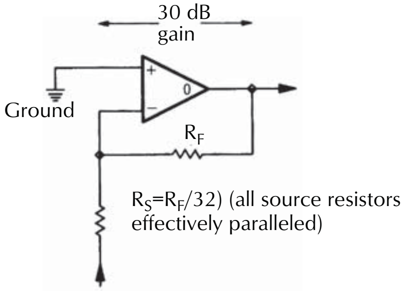
25.8.4.4 Ground Noise in Virtual-Earth Mixers

A virtual-earth mix-amp unavoidably amplifies ground noise. Fig. 25-28A tells the story. For instance, a multitrack mix-amp can typically have 32 sources applied to it; the through gain from any source is unity (assuming the source resistors equal the feedback resistor), but the real electronic gain of the circuit is 33 or a touch over 30 dB. Redrawing the circuit slightly in Fig. 25-28B shows exactly what this 30 dB is amplifying. Consider as a clue that which is directly applied to the noninverting input of the op-amp—the ground! True, it is amplifying the noise due to the resistors and the internal noise mechanisms of the device, but for our argument here, it is amplifying ground. In any reasonably sized console, providing no sources are grossly out of proportion to the majority, ground noise is pretty random and noisy in character. The result is that, on being amplified up, it serves to make the mix-amp apparently much noisier than would be expected from calculation. In suspect systems it has been found to be the predominant noise source. It is no accident that the real electronic gain of a mix-amp is also known as its *noise gain*.

It is truly astonishing what attention to virtual-earth mixer grounding can have on bus noise figures. For mix-amps, practical noise performance has little to do with the device employed and nearly everything to do with grounding.



A. Conventional mix-amp.



B. Circuit redrawn as a noninverting amplifier with ground as signal (noise) source.

Figure 25-28. Virtual-earth mix amplifier as amplifier of ground-borne noise.

25.8.4.5 Reactive Ground Effects

Noise generation due to grounds is not limited to the resistance predominant in the ground wiring at audio frequencies. At radio frequencies (well within the bandwidths of modern op-amps) even fairly short ground wires and buses can have very significant reactances, dramatically raising the effective ground impedance. This not so much reduces the isolation between the various stages as directly couples them together. All the inherent RF noise and instabilities of the stages become intermodulated (by the nonlinearity of the device at those frequencies) to make their presence felt as yet more audible and measurable noise.

A good "shock horror" extreme example, though described in simplistic theoretical terms, manifests itself sometimes dramatically in practice and can be called the *standing on one leg effect.*

The box in Fig. 25-29 represents a device that relies on a wire to be connected to the ground mass. It looks all right, and so it is, apart from the fact that at certain radio frequencies the wire is electrically ¼ wavelength or an odd multiple of ¼ wavelength. In accordance with transmission-line theory our innocuous bit of wire turns into a tuned line transforming the zero impedance of the ground to an infinite impedance at the other end. The result is that the device is totally decoupled from ground at those frequencies. Practical consequences of this, of course, vary, from instability at very high frequencies on cards with long supply and ground leads to painful, unreasonable susceptibility to RF in otherwise wholesome items of equipment.

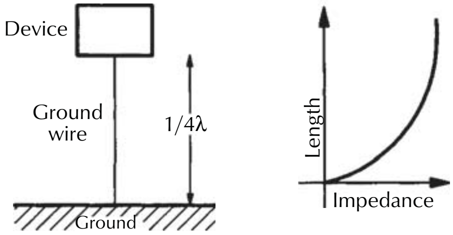


Figure 25-29. Standing-on-one-leg effect.

25.9 Signal Switching and Routing

Signal routing within the channel and other areas of the system is a touchy affair that has always been an area of much discontent for console designers, especially since the advent of in-line consoles and remotable and assignable systems. There are always standard relays, but these have lost, justifiably, a lot of appeal in the light of ac technologies.

25.9.1 Relays

Unless they are of the expensive miniature IC package variety, relays tend to be big, heavy, eventually unreliable, mechanically noisy, and a nuisance to implement electronically. They also demand support circuitry such as back-emf protection diodes and drive transistors for a realistically operable system. The coils, being inductive in nature, draw a surprisingly large instantaneous on current and release an equally surprisingly large amount of back-emf energy when deactivated. Both of these—through mutual-inductance coupling, dubious common ground paths (even as far back as the master ground termination in separated supply systems), soft power supplies, and even mechanical microphonic effects—tend to impinge themselves on audio signal paths as clicks, splats, pings, and other assorted bumps. Of course, it's possible to have silent relay switching. However, after designing in separate ground unrelated power supplies of considerable heft, spatially separating the relays from the audio (preferably on another card), working out the drive interfaces, and liberally sprinkling the whole issue with diodes, resistors, and capacitors to tame the spiky transients, the circuit becomes very complicated.

Certain routing applications do implicitly require relays and their lack of concern about the amount of dc and either common-mode or differential signals of absurd quantities that may accompany the audio in balanced networks. Such circumstances are to be found anywhere a telephone line is used.

This is almost specifically a broadcaster’s concern, where many external high-quality sources appear down phone lines and need to be routed before hitting either the internal distribution amplifier system of the station or perhaps even a console line input directly. Outside source selection, as it’s called, does not fortunately have the same splat-elimination constraints as intraconsole switching, since the signal is nearly always of high level, balanced, and riding with at least a little dc (which will unavoidably click upon switching); most importantly the selector is very unlikely to be switched while actually live on air.

25.9.2 Electronic Switching

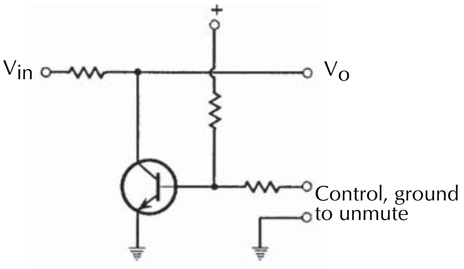
The wish list for an audio switch is simple:

1. It has an infinite off impedance.
2. It has a zero on impedance.
3. It has a control signal that is isolated from and does not impinge on the through signal path.
4. It costs nothing.

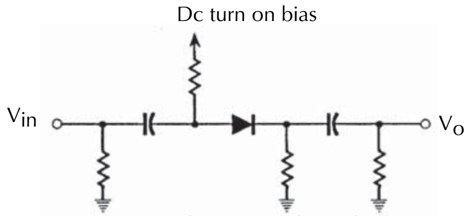
In the real world, of course, some leeway has to be given, but, fortunately, the tradeoffs are more in subtleties than in these basics.

Transistors are out of the picture right away despite their high on-off impedance ratios, because they are essentially unidirectional in current flow, and the control port (the base) is actually half of the signal path as well. In certain circumstances they have been used in the place of relays as a soft output muting clamp as in Fig. 25-30A.

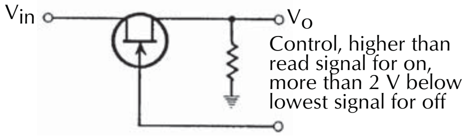
Diodes, Fig. 25-30B, are used extensively for signal routing in RF equipment, with the required signal riding on a relatively large dc bias that overcomes the diode forward voltage drop making it a low-impedance path, and correspondingly turned off by a large back bias. Considering that in some audio design cases getting rid of a handful of microvolts dc can be an ordeal, somehow a couple of dozen volts hurling about lacks a certain appeal. Typically, switching diodes for RF with a PIN structure are chosen; their very small capacitances are considerably less parametric with respect to varying reverse bias, so minimizing automodulation and consequent distortions.



A. Transistor as a signal switch.



B. Diode as a signal switch.


V_{in}
V_O
Control, higher than read signal for on, more than 2 V below lowest signal for off

C. Field effect transistor as a signal switch.

Figure 25-30. Solid state devices as simple switches.

FETs have been and still are used extensively for signal switching. They again have a high on-off ratio, and the control port (the gate) is of extremely high impedance and well isolated from the signal path, but the gate on-off voltage levels are a bit awkward for interfacing with logic control signals. They also define signal head room through the switch, based on the gate on-off biasing voltage range. It is bidirectional, its channel path being essentially just a voltage-controlled resistor, but the on resistance tends to vary with the varying audio voltage across it (auto modulation); distortion in the more basic FET switching configurations can be a problem. However, they are workable, Fig. 25-30C.

25.9.3 MOSFETs and CMOS

Closely related to FETs are metal-oxide-semiconductor field effect transistors (MOSFETs). They have a different chemical structure and physical construction but have essentially similar characteristics with the exceptions that the gate is of even higher impedance, and the control voltage swing required is easier to deal with. Complementary MOSFET (CMOS) elements, connected back to back to form close to ideal bidirectional analog transmission gates, are manufactured in all manner of variations and packages by IC manufacturers. At extremes of performance minor control-port breakthrough (charge injection) can rear its head and be taken into consideration.

Early versions of CMOS transmission gates had some rather untoward vices. They were raw CMOS elements, and one of their main attributes, the extremely high impedances in their off states and of their control ports, made them liable to destruction by normal amounts of static electricity. Also, they tended to latch up easily if any of the MOS junctions inadvertently got reverse biased into conduction (this happened easily if the signal voltage passing through a gate even momentarily exceeded the supply voltage). Most present devices are now gate protected to prevent static blatting, and the worst that happens with the audio signal exceeding the switch supply voltage by a small amount is that the switch breaks over (i.e., conducts audio momentarily). It does not result in the fatal consequences it once did.

Perhaps the best-known and most-used switch of this kind is the 4016 (and its younger brother the 4066, which is essentially identical but for a lower on resistance). It is a 14-pin dual package containing four independently controllable CMOS transmission gates. Each gate can pass up to the IC's supply voltage (typically 18 Vdc) into a load exceeding 10 kΩ with a distortion of about 0.4% in rudimentary switching formats. Obviously, both the distortion figure and the head room availability of 18 dB above 0.775 V (for an 18 Vdc supply) are both woefully inadequate by today's expected console standards. Another less obvious pitfall is the decreasing switch isolation at high frequencies due to leakage capacitance across the gate.

Fig. 25-31A gives a typical representation of the variation of the on resistance of a CMOS transmission gate with signal voltage applied to the gate. This variation in resistance is, of course, the source of the distortion. If we could restrict the signal voltage to within that (linear) bit in the middle, or better still virtually eliminate the signal voltage altogether, our problem would go away.

Placing the switching element right up against a virtual-earth ground point, as in Fig. 25-32A, achieves this signal voltage elimination; the switch now behaves as a two-state resistor. When closed, the on resistance variation, which will be small anyway because of the very low voltage swing across it, will be effectively swamped by the (relatively) much larger series resistance. When open, the off resistance extends the total series resistance to a value approaching infinity. In practice, the on-off ratio is not really adequate. Capacitance across printed circuit tracks and in the device encapsulation itself, combined with common-ground current and other essentially flat-response crosstalk mechanisms, results in a cross-switch leakage characteristic ultimately rising 6 dB/octave against frequency. Also, despite the fact that the distortion problem is now largely resolved, there still remains a head room problem when the switch is open. If the source voltage presented to the series resistor exceeds that of the power supply of the CMOS gates, the gate will break over, turning on for that excessive portion of the input waveform.

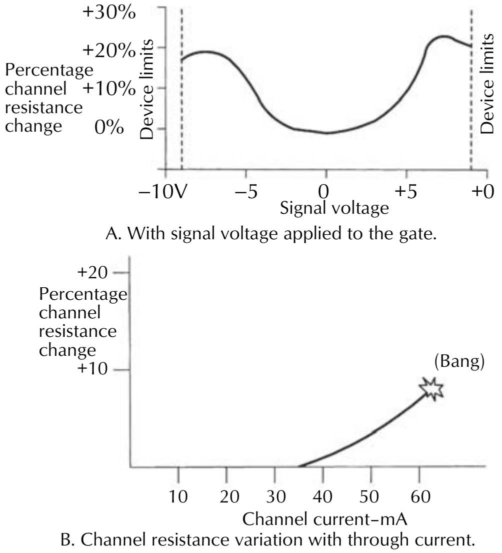
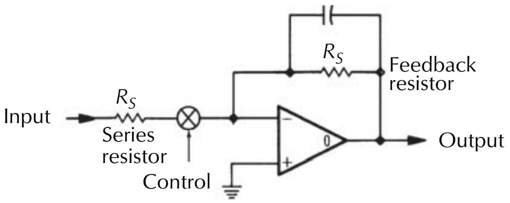


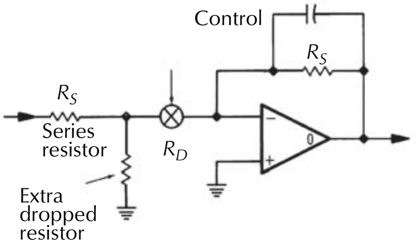
Figure 25-31. Typical CMOS transmission gate linearity.

Attenuating the source signal by the needed amount before it hits the gate skirts this hangup. Unfortunately, this worsens the noise gain of the virtual-ground amplifier by the amount of that attenuation. In Fig. 25-32B dropping an equal-value resistor to the series resistor to ground from its junction with the gate is a working approach. The maximum signal that can be present across the gate when off is now half that previously, which is usually more than enough attenuation to prevent breakover. This 6 dB loss is magically made up for in the on mode because the source resistance of the signal into the amplifier is now halved (series resistance effectively in parallel with the dropped resistor).

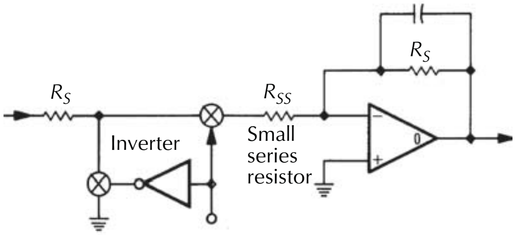
Incidentally, the crosstalk improves as a consequence by almost 6 dB—less signal voltage actually within the chip. For many practical purposes, this switching configuration, with its performance limitations as defined, is quite adequate. For instance, the



A. With signal voltage applied against a virtual-earth point.



B. With dropping to ground an equal-value resistor to the series resistor.



C. With second analog transmission gate replacing the dropped resistor.

Figure 25-32. Switching arrangements using CMOS transmission gates.

noise and crosstalk characteristics are a good order of magnitude superior to any analog multitrack recorder, so this element can be a good choice for an inexpensive track assignment routing matrix.

25.9.4 Potentiometric Switching

A refinement of this element—in fact, really an extension of the same principle—is shown in Fig. 25-32C. Here, a second analog transmission gate replaces the dropped resistor and is driven through an inverter from the control line for the original gate, arranging for it to be on when the other is off and vice versa. When the original gate is on, there is very little potential across either of the gates (they're both at virtual ground from the op-amp). Similarly, there is little potential across either of the gates when the second gate is on, since it is tying the series resistor to ground and the open gate is between ground and virtual ground. Crosstalk is dramatically improved when the element is off because any signal present at the series resistor faces the double attenuation of the series resistor tied to ground by the on second gate followed by the off original gate into the virtual-earth input of the op-amp. In the on mode of the element, there is no input attenuation; hence, there is no gain and no extra noise contribution from the amplifier. The only limitation now to the cross-switch leakage characteristic of this switching element is printed-circuit card layout and grounding arrangements. Given a good home, this element is virtually unmeasurable.

It does, however, have one quirk that may preclude its use in some places. Unless a great deal of care is used to arrange complementary on-off switching timing for the two gates, they are both momentarily partially on together during a switching transition. This, for an instant, ties the virtual-earth amp input to ground via the quite low half-on impedances of the two series gates, creating an instantaneous burst of extremely high gain from the amp; this shows as a transient of noise or worse still as a splat if any dc offset is present at the virtual-earth point. It can be minimized, or at least the extent of the transient defined, by a small value resistor (R_{ss}) in series with the input, Fig. 25-32C. This will, of course, increase the signal voltage across the gates and increase the distortion, so a compromise has to be struck to suit the given application. Even so, excessive distortion owing to this has never shown itself to be a problem.

25.9.5 Minimizing Noise

To reduce the thermal noise contribution as part of the circuit noise performance, the resistances involved in switching should be as low as practically possible consistent with device limitations and the ground current arrangements. The feedback resistor around the virtual-earth stage is limited by the output drive capability of the op-amp, bearing in mind it has to drive its load, too. Fig. 25-31B demonstrates a typical channel resistance variation of a CMOS switching element with through current. It behaves linearly until about 40 mA, which actually compares more than favorably with the output drive current capability of an op-amp. (FETs are excellent constant-current sources, self-limiting in nature.) As a rule of thumb then, the resistors used around analog gate switching circuits can be as low as 2.2 kΩ without exceeding device limitations; the high-output current capability of the 5534 can be used to good effect here if the drive for lowest possible thermal noise is that important. Generally, ground-borne noise generously provides a noise floor well before this theoretical limit is attained; the whole lowest-impedance question becomes self-defeating eventually. The more current is chucked around, the worse the ground noise is going to become.

25.9.6 Practical Matrix Example

The 4000 series of CMOS devices, which are very commonly used, have one important feature at odds with general mixer technology—their maximum supply voltages. The earlier 4000A series were limited to a 15 Vdc total (as compared to the 30 Vdc or 36 Vdc total commonly used in console design), while the more recent buffered B series can stand 18 Vdc. More recent families (HC, e.g.) nearly always adhere to the common digital electronics supply of 5 V (actually 7 V max). An advantage of the earlier series was that a separately regulated 5 V supply wasn't necessary. Nowadays, though, there is nearly always 5 V running around for this, that, or the other and it is not a difficulty to create a sufficiently well-regulated and quiet 5 V supply for HC switches. Given the virtual-earth switching technique described, this diminution of supply is immaterial.

Most CMOS families have a wide range of switch configurations available; some specialist devices integrate quite large arrays. By way of example, Fig. 25-33 shows one mixer channel's worth of a digitally assigned 32-track routing matrix, designed around a pair of Harris HI506A 16-way multiplexers. This part, for example, contains 16 analog transmission gates tied to one common output (which we will rename *input*). Each of the free ends of the gates ties directly to a mix bus. They all share a common series source resistor via the input port. Since only one of these gates can be open at a time (the one corresponding to the binary 4-bit address code on the address inputs), there is no possibility of two or more buses being inadvertently shorted. The device manufacturers proudly point out the break-before-make delay in switching, meaning that a newly selected gate waits until the previous one has delatched, so there is no momentary switching short.

25.9.7 Matrix Crosstalk

Crosstalk with this configuration, which you will notice is a variation between Figs. 25-32A and C, is extremely good. Again, there is the double attenuation of the series resistor via an on-gate to a virtual-earth bus (some 20 dB isolation to start with), followed by the internal isolation between buses owing to the off-gate impedances into all the other virtually zero-impedance mixing buses (some additional 70–80 dB). A slightly more critical crosstalk situation could exist when all the gates are turned off (by tying the HI506A enable low) since the first set of attenuation no longer exists; the switches' common point would no longer be tied to a virtual earth by any of the elements. This is why external switching elements (IC3a and IC3b) are arranged to tie the end junction of the series resistor and the HI506A input common point to ground whenever the enable lines are low.

Crosstalk is now completely down to the interconnections to this card, power supply decoupling, solid and correct ground paths, but mostly inductive and bus/earth/bus eddy-current coupling between the virtual-earth buses themselves. This is yet another design area where performance is completely determined by mechanical considerations.

25.9.8 16-Track or 32-Track Routing

The switching card described above may be configured merely by changing two wire links in two different routing formats. The first enables a stereo pair of signals (i.e., the panned outputs of a channel) to be routed to adjacent odd/even pairs of outputs (i.e., 1 and 2, 7 and 8, 27 and 28, and so on), where the odd numbers represent left and the even numbers represent right. Either odds or evens may be accessed singly by suitable feeds to the odds enable and evens enable control inputs. Quite obviously these also facilitate disabling (turning off completely) the routing.

A 4-bit binary control bus selects which pair of the possible 16 pairs may be accessed, so these six control lines are all that need to be extended to the channel module where simple switching performs all routing requirements.

When the aforementioned wiring links are made in the fashion shown in Fig. 25-33, the card becomes configured as a 1-source-into-32 destination switcher, necessitating some control function changes. Evens enable becomes the additional highest significant bit of the destination address code (5 bits are needed for 32 combinations), while odds enable turns into the enable/disable control of the switcher. (The benefit, in both modes, of disabling the switcher when not actually in use is that it removes the feed totally from the destination buses. Their performance is not impaired at all, and a preselected routing setup on the address lines is not disturbed.) With the same signal applied to both the audio inputs, it is now possible to access any one of the 32 buses singly.

25.9.9 Processor Control

The seemingly great mass of logic circuitry enclosed in the dotted lines allows the card to be controlled by a computer or microcontroller's I/O lines (depending on the scale of the system). Very little additional decoding is necessary to allow this interface to hang straight onto a CPU bus. It's really just six flip-flops acting as memory elements (so that the card can remember what the CPU has told it to do) and six tristate buffers that, on request, tell the CPU what the card is actually doing. This memory both saves the CPU from having to store the matrix routing information somewhere else and also acts as a very useful diagnostic aid to help find out what isn't doing what, where, and why.

For ordinary direct operation, this logic can be left off the card completely and linked across (between the Xs on the diagram). The NAND gates in the top left-hand corner merely organize the CPU bus information to fire the appropriate clock, enables, and resets to the memory elements.

CMOS 4000 series logic operating at 5 Vdc is not the fastest logic family in existence and is too slow for most microprocessor CMOS to drive directly. The circuitry shown here will operate successfully from a microcontroller running at a MHz or two, which is very slow in the computer world. This slowness is not a problem in reality, since the practical way of dealing with this is to hang the entire switching matrix logic system off a bunch of the CPU input/output ports, masquerading as a local address/data/control bus system at a fairly leisurely software-controlled rate.

A convenient 16 input-output (I/O) line is required (two lots of eight, handy for PCs). A single 8255 or 6850 Peripheral Interface Adapter (PIA) would handle this matrix. Being software controlled, the I/O lines may be timed a little more gently than the hardware-determined processor buses.

A separate address decoder card however takes many of the card address bits that are required (5 for 32, 6 for 64, 7 for 128) and generates the decoded feeds for the card enable (CE) on each matrix card. This is very simply accomplished with a daisy chain of 4028 binary-to-decimal decoders, Fig. 25-34.

This slow interface has the single benefit that it is fully featured; it is common with faster and more capable processors with plenty of cheap memory aboard to not require the readback facility; nearly all the interface logic can be replaced by parallel bus latches such as HC373s with suitable address decoding. HC of course will operate at reasonably high clock rates, even directly off a processor's data buses. FPGAs, the ultimate octopus of peripheral handling, can handily absorb all the described interfacing.

25.9.10 Audio Path

In the bottom left-hand corner of Fig. 25-33 is an analog mix-amp and line amp, which are the group output stages for the channel to which the particular matrix card is relevant and they are as close as they can get to the buses.

The mix-bus input is tied on the back of the edge connector in the card frame to the bus it is responsible for sensing. This ensures card replaceability and redundancy; individually doctored cards are the kiss of death from a maintenance standpoint since there is no means of getting a given path going again (in the event of a failure) without actually fixing the fault. The old standby of swapping cards wouldn't work; it is always best to keep individualism off the cards.

Note that no values are attributed to the feedback capacitor around the mix-amp, since this not only has to compensate for the amplifier's own tendency to instability but for the added irritation to this of the bus impedance—an unknown until actual construction. Also note a capacitor across part of the switcher input series resistance. This provides a variable high-frequency kick, which can be of assistance in sorting out frequency and phase response quirks in problematic bus systems. This is, fortunately, very rarely needed and is provided just in case.

Fig. 25-35 shows the audio path through the switcher, devoid of frills. The 1 kΩ resistor RS, which does not appear in Fig. 25-35, is internal to the HI506A, appearing on each of the switch inputs. Although a minor nuisance in this application, which means the CMOS switches are not actually switching a zero impedance, they are part of the device's internal protection against, principally, static electricity damage—a worthwhile sacrifice.

The total source impedance before the bus is about 9.9 kΩ, which with the addition of the 100 Ω buffering resistor becomes 10 kΩ before the virtual-earth input of the mix-amp. A 4.7 kΩ gain-trim preset in series with 8.2 kΩ gives a gain determining feedback resistor swing of approximately 8.2–12.9 kΩ, which corresponds to a swing of −1.7 dB to +2.2 dB.

The line amp is a simple beefed-up inverting amplifier necessary to maintain the absolute input-output phase relationship.

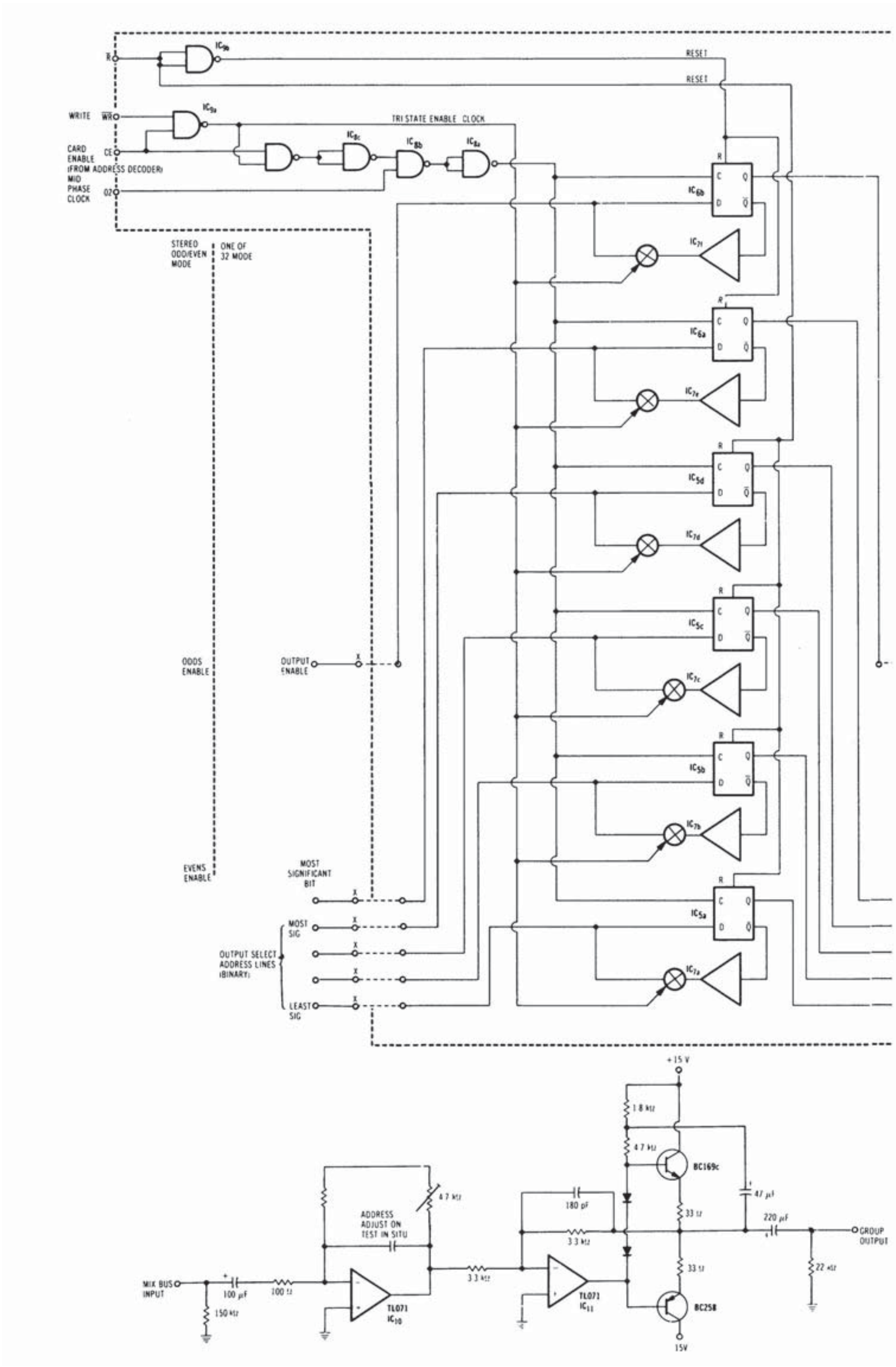

R̄
IC9b
RESET
RESET
WRITE
W̄R̄
IC9a
TRI STATE ENABLE CLOCK
CARD ENABLE (FROM ADDRESS DECODER)
CE
MID PHASE CLOCK
02
IC8c
IC8b
IC8a
STEREO ODD/EVEN MODE
ONE OF 32 MODE
ODDS ENABLE
EVENS ENABLE
OUTPUT ENABLE
MOST SIGNIFICANT BIT
MOST SIG
OUTPUT SELECT ADDRESS LINES (BINARY)
LEAST SIG
X
R
C
Q
D
Q̄
IC6b
IC7f
IC6a
IC7e
IC5d
IC7d
IC5c
IC7c
IC5b
IC7b
IC5a
IC7a
+15 V
1.8 kΩ
4.7 kΩ
BC169c
47 μF
4.7 kΩ
ADDRESS ADJUST ON TEST IN SITU
180 pF
33 Ω
3.3 kΩ
220 μF
GROUP OUTPUT
33 Ω
22 kΩ
MIX BUS INPUT
100 μF
100 Ω
150 kΩ
3.3 kΩ
TL071 IC10
TL071 IC11
BC258
15V

Figure 25-33. Processor-controllable matrix routing circuit.

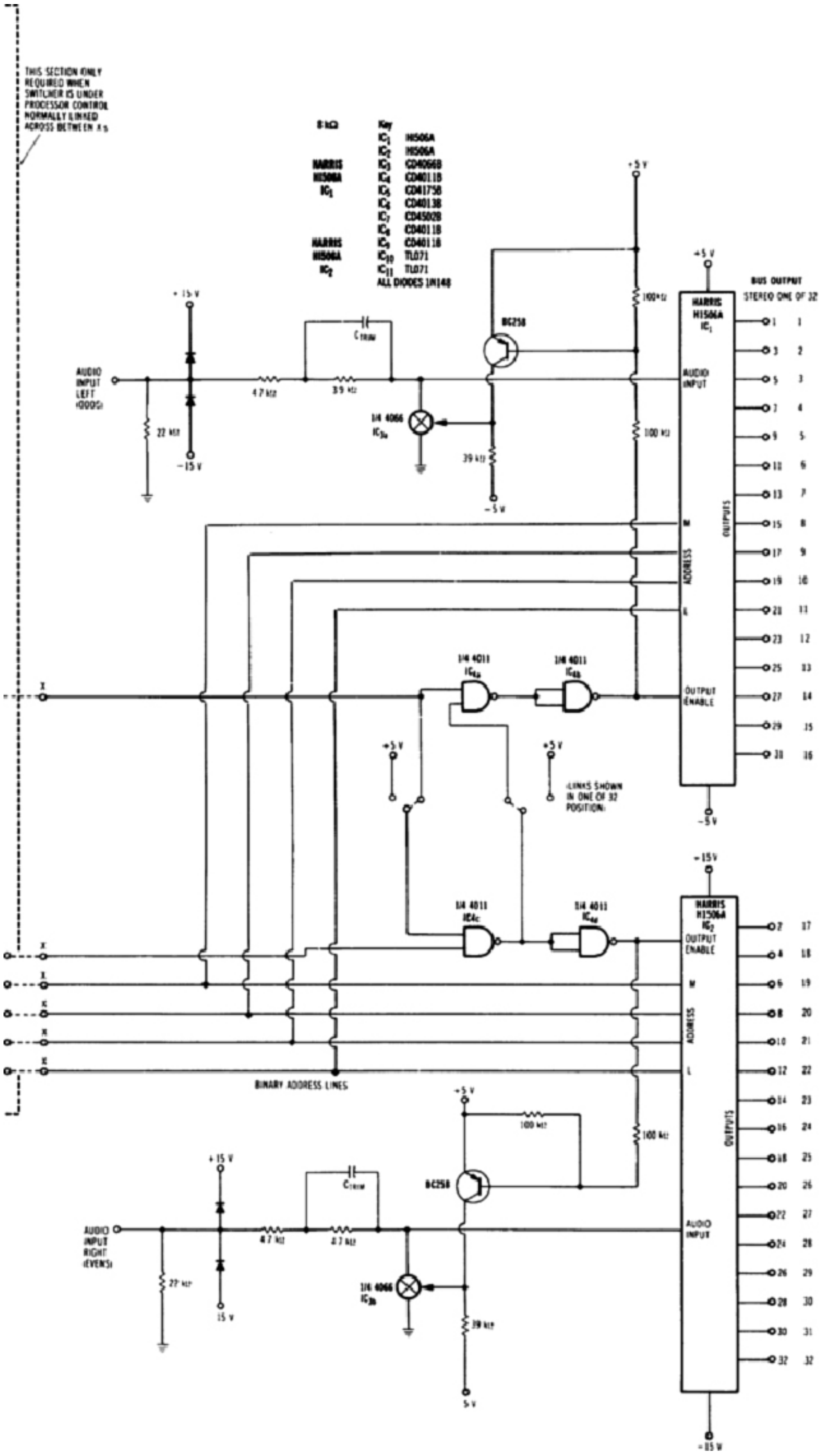

THIS SECTION ONLY REQUIRED WHEN SWITCHER IS UNDER PROCESSOR CONTROL NORMALLY LINKED ACROSS BETWEEN X's
8 kΩ
Key
IC1 HI506A
IC2 HI506A
IC3 CD4066B
IC4 CD4011B
IC5 CD4175B
IC6 CD4013B
IC7 CD4502B
IC8 CD4011B
IC9 CD4011B
IC10 TL071
IC11 TL071
ALL DIODES 1N148
HARRIS HI506A IC1
HARRIS HI506A IC2
AUDIO INPUT LEFT (ODDS)
AUDIO INPUT RIGHT (EVENS)
+15 V
−15 V
22 kΩ
27 kΩ
4.7 kΩ
3.9 kΩ
C TRIM
1/4 4066 IC3a
1/4 4066 IC3b
39 kΩ
−5 V
+5 V
BC258
100 kΩ
1/4 4011 IC4a
1/4 4011 IC4b
1/4 4011 IC4c
1/4 4011 IC4d
LINKS SHOWN IN ONE OF 32 POSITION
BINARY ADDRESS LINES
ADDRESS
AUDIO INPUT
OUTPUT ENABLE
OUTPUTS
BUS OUTPUT STEREO ONE OF 32
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31
2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
+15V
−15V

Figure 25-33. Processor-controllable matrix routing circuit. Continued.

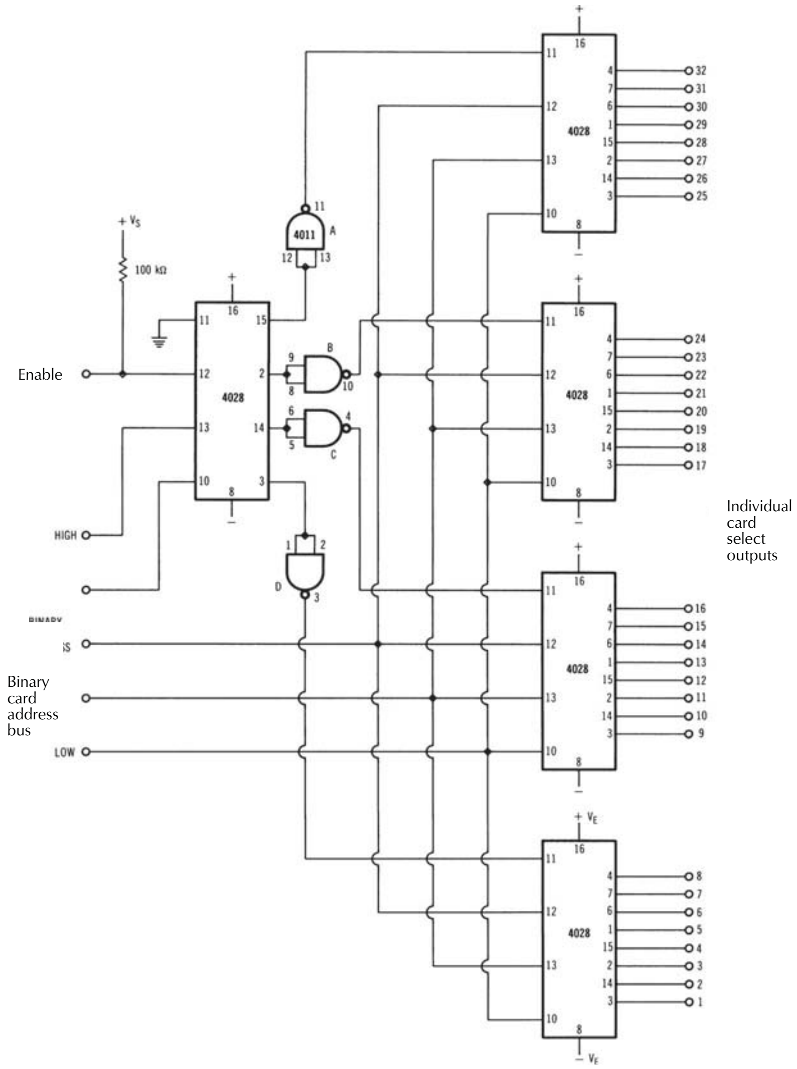

+ V_S
100 kΩ
Enable
HIGH
LOW
Binary card address bus
4011
A
B
C
D
4028
4028
4028
4028
4028
+ V_E
− V_E
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Individual card select outputs

Figure 25-34. Matrix channel-decoding logic.

25.10 Input Amplifier Design

A console is expected to accept signals of wildly varying input level and impedance while producing a uniformly consistent output capable of being deposited in the tightly defined container that is a recording track, or similarly defined output.

Fortunately, industry standards provide at least some clues as to what mixers are likely to have applied to them. Nevertheless, these standards can obviously do nothing to alter the physics of the operation of the assorted transducers and sources in common use; the disparity in the treatment required between a dynamic microphone and a recorder output totally precludes a universal input stage.

Mixer front-end design tends to be a little like working on a grown-up jigsaw puzzle where all the important pieces perversely refuse to fit. It's delightful to discover or cultivate some that fit nicely, like in line-level input stages. This euphoria is chipped away by the problems inherent in other areas, notably microphone input stages.

Optimizing input noise performance in a dynamic microphone preamp is quite an operation, juggling a

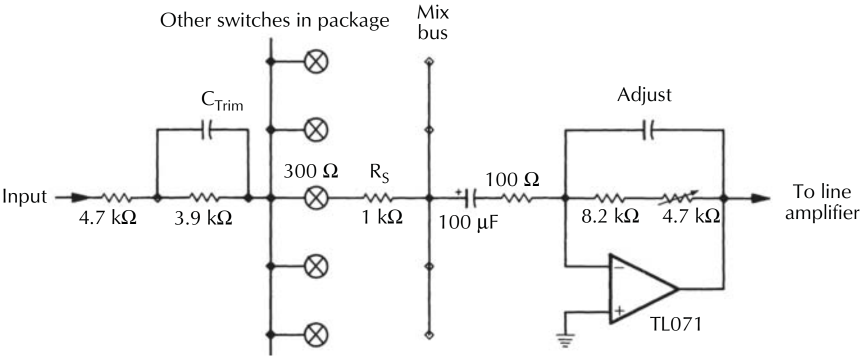


Figure 25-35. Switcher audio path.

seemingly endless number of variables. A dynamic microphone may be represented (a little simplistically) as a voltage source in series with a fairly lossy inductance representing a midband impedance typically of 150–300 Ω, Fig. 25-36. Being a transducer and, of necessity, mechanical in nature, many complex varying motional impedance effects contribute to the overall scene, as do the effects of matching transformers used in many microphones. For most design purposes, however, this simplistic electrical analog can suffice. The low impedance commonly and conventionally used is primarily to mitigate high-frequency attenuation effects due to inevitable cable capacitance. Despite the fact that the characteristic impedance of microphone cable is not too far removed from that of our typical sources, the runs are so short in wavelength terms that transmission-line "think" is not really applicable and the cable just looks like a distributed capacitance. This, in practical circumstances, amounts to a large value of capacitance that the transducer must drive along with its load. Unfortunately, the impedance is not low enough that it may be treated as a pure voltage source; therefore a tiny signal at a finite impedance must be ferreted out and treated with care for optimum performance.

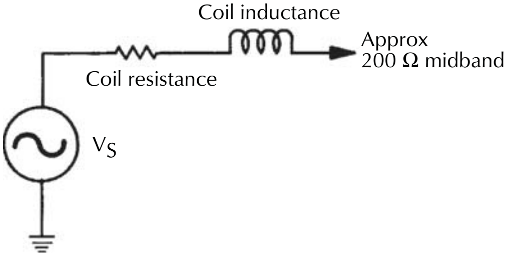


Figure 25-36. Simplistic dynamic microphone model.

25.10.1 Power Transfer and Termination Impedance

Textbooks on electrical theory state that to extract maximum power from a given source the optimum load is equal in value to the source impedance. In the dynamic microphone, it is of doubtful (if any) value. We've squeezed all the energy possible from the generator, but to what end? Given that most electronic amplifiers of the type useful in low-noise applications are of relatively high-input impedance (i.e., voltage amplifiers), then the terminating resistance that largely defines the load of the microphone would, in fact, dissipate most of our hard-won power. It is the output voltage capability of the source that is of greatest use here, not the power. So, as can be seen in Fig. 25-37, matching source and load impedances does a very effective job of sacrificing 6 dB of signal level that has to be made up in the succeeding amplifier. This does not imply that the noise performance is 6 dB worse than possible, since the source impedance as seen by the (assumedly perfect) amplifier is now a parallel of the microphone and its matching load; hence, it is about half the value of either. The thermal noise generation of this combined source is consequently 3 dB less; so although the voltage is down 6 dB the noise performance is only degraded 3 dB by such a termination. Still, it is better not to throw away a good 3 dB before even starting to hassle with the amplifier itself.

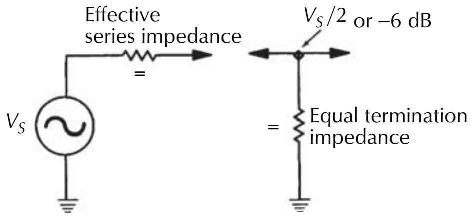


Figure 25-37. Matching—how to lose 6 dB.

Another good reason for not terminating with an equal or any fairly low resistance is the effect on microphone response and subjective quality. Having an inductive characteristic, the dynamic microphone capsule has an impedance that steadily rises with frequency, becoming predominant at high audio frequencies where the inductive reactance of the source is large with respect to the coil-winding resistance. When terminated with a relatively low resistance, the complex impedance of the capsule and the termination resistor form a single-order 6 dB/octave low-pass filter, gracefully rolling off the high-frequency output of the microphone. Not too useful.

With a fairly hefty cable capacitance, the system is no longer graceful; the complete network now looks like a rather rough second-order filter. There isn't too much to be done about that; regardless of termination method, cable capacitance is here to stay and is always a consideration unless the preamp is remoted to or close to the microphone itself.

25.10.2 Optimizing Noise Performance

Amplifiers are not perfect. For noise criteria, the first device that the signal hits in the amplifier is the key one, since the noise it generates usually masks—by a large margin—noise from all succeeding stages.

All practical amplifying devices are subject to a variety of internal noise-generating mechanisms, including thermal noise generation. When measured, these give rise to some important values; namely, the input noise voltage, the input noise current, and the ratio between those two that is in effect the input noise impedance. This becomes all important in a little while. For the most part, bipolar transistors—either standard or more usually large-geometry and sometimes multiparalleled—are used as front-end devices both in discrete designs and op-amp IC packages in this application so much of the following relates to both packages.

These noise voltages and currents alter in both individual magnitude and ratio to each other with differing electrical parameters, especially collector current. Predictably, as this current decreases, so does the noise current (most of the noise is due to minor random discontinuities in device currents); the ratio between the noise voltage and current—or noise impedance—may be altered in this fashion.

Thermal noise generation is common to all resistive elements. The amount is related to both the temperature and the bandwidth across which it is measured; an increase in either will increase proportionally the noise power generated. Under identical circumstances, the noise power that is generated by any values of resistance is the same. Differing resistor values merely serve to create differing ratios of noise voltage and noise current; the product of the two always equals the same noise power. This particular noise phenomenon, thermal noise or Johnson noise, is totally unavoidable because the nature of atomic structure is such that when things get hot and bothered, they grind and shuffle about randomly, creating electrical disturbances white in spectra (i.e., equal energy per cycle bandwidth).

Even the real (resistive) part of the complex impedance of a dynamic microphone generates thermal noise; this ensures that there is a rigidly defined minimum noise value that cannot be improved upon.

25.10.3 Noise Figure

The difference between the noise floor defined by thermal noise and the measured noise value of a practical system is known as the noise figure (NF) and is measured in decibels (Noise Figure = System Noise – Theoretical Noise). The noise output from a resistor or the real part of an impedance is calculable and predictable—Herr Boltzmann rules. A direct comparison of the noise voltage measured at the output of an amplifier due to a resistor applied to the amplifier input versus the noise voltage expected of the resistor on its own is possible just by simply subtracting the measured gain of the amplifier. This is a measure of NF.

An interesting effect occurs when, with any given set of electrical parameters set up for the amplifier front-end device, the source resistance is steadily changed in value. A distinct dip in the NF occurs, Fig. 25-38, and the value of the resistor at which this dip occurs changes as the device parameters are changed (collector current primarily). For the usually predominant noise mechanism (thermal noise), a minimum NF occurs with a tiny amount of collector current (5–50 μA) and a high source resistance (50 kΩ up). Without diving into the mathematics, the nulling is a balancing of interaction between the external noise source and the internal voltage and current noise generators.

25.10.4 Inverse-Frequency Noise

There is another major noise mechanism inherent to semiconductors. It is the low-frequency (inverse level with respect to frequency) noise—a burbly, bumping-type noise caused by the semiconductor surface generating and recombining sporadic currents—most prevalent in dirty devices but present to a degree in all. It is subjectively apparent and has to be

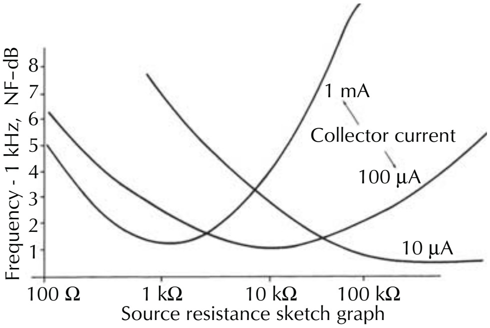


Figure 25-38. Bipolar noise curves (noise figure curves for a good pnp front-end transistor for collector current versus source resistance).

considered. Measured alone, low-frequency noise has its own set of collector current and source resistance nulls, usually far higher in current and lower in resistance than for thermal noise.

Commonly known as *1/F noise*—implying its predominance at very low frequencies—it is often specified by way of the frequency at which it is contributing the same amount of noise as the device's effective thermal noise. Below this knee frequency 1/F noise predominates. A good clean device will have a knee frequency below 10 Hz; judicious filtering along the signal path can render 1/F noise unimportant within a system, but it remains a serious consideration within each individual amplifier stage.

25.10.5 Optimum Source Impedance (OSI)

A compromise has to be struck. To make a generalization, a 100 µA collector current and a 10 kΩ source impedance for a typical low-noise pnp transistor seem about right. (Pnp transistors are commonly used in this area due to slightly better low-frequency performance figures over npn types.) The source resistance value is that at which the device is optimally quiet for audio purposes and is known as the optimum source impedance (OSI). Incidentally, this impedance has absolutely nothing to do with the kind of circuit configuration in which the device may be. Whether it be in a common-base amplifier with an input impedance of 50 Ω or in a totem-pole front end with bootstrapping and a consequential input impedance of over 10 MΩ, it doesn't matter. The source impedance for optimum noise performance stays at 10 kΩ, or whatever, provided that the collector current is the same in all cases. Optimum source impedance has nothing to do with input impedance.

This optimum impedance varies depending on the type of input device used. For an FET, the noise figure typically obtainable drops to an amazingly low value but, unfortunately, at a substantially useless impedance of several dozen megohms. Even supposing it were practical to provide a source impedance of that magnitude, the whole arrangement would be so sensitive to any electromagnetic fields (such as RF) that even tiny amounts present would obliterate the noise advantage. The design and construction of capacitor microphones using FET front-ends highlight the hazards; the end results often show such capacitor microphones can be several dB noisier than a dynamic microphone/front-end combination.

Good bipolar transistors have OSIs in the region of 5 kΩ to 15 kΩ, whether discrete or as part of an IC amplifier package. Fortunately, these values closely coincide with the source resistance value that provides for optimum flatness of device transfer characteristics. This helps a long way toward best frequency versus phase linearity, which translates to enhanced stability in a typical high negative-feedback amplifier configuration.

Fig. 25-39 shows the effect of altering the source impedance into such an amplifier (using a conventional bipolar transistor input device) on output frequency response. The droop is due to the excessively high source impedance reacting against the device base-emitter, board, and wiring capacitances to form a low-pass filter. The high-frequency kink is a practical effect of the curious mechanism; when a bipolar transistor is fed from an impedance approaching zero, its high-frequency gain-bandwidth characteristic extends dramatically, radically altering the phase margin and, consequently, the stability of an amp designed and compensated for more ordinary operating circumstances. The kink is a resonance within the amplifier loop caused by erosion of phase margin resulting from this mechanism. It is but an uncomfortably short step from oscillation.

As can be seen from the graph in Fig. 25-39, the response is maximally flat at a source resistance of around 10 kΩ, about the same value as the OSI for optimum noise performance for the same configuration. A problem to reconcile is that our practical source impedance is nominally 200 Ω for a dynamic microphone, whereas the OSI for the best conventional input devices is around 10 kΩ. How do we make the two fit?

25.10.6 Microphone Transformers

There are some horrible stories about how bad transformers are. Properly designed and applied, however,

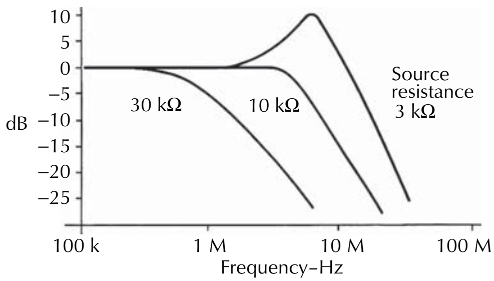


Figure 25-39. Source impedance versus bandwidth gain versus frequency for a typical follower-connected operational amplifier highlighting effects on response of source impedance on input device.

they do offer a good to excellent solution to impedance matching and sundry other problems facing input stage design. Simplistically, a transformer is a magnetically soft core around which are two windings, the voltage ratio between the two being equal to the ratio of the number of turns on each. The impedance ratio is the turns ratio squared (e.g., a 10:1 turns ratio corresponds to a 100:1 impedance ratio) because power output cannot exceed power input. If the voltage is stepped up ten times, the output current must be stepped down ten times. Impedance, which is the ratio of voltage to current, is consequently the square of the transformed voltage or current ratio, see Chapter 11.

Given this, it is a simple matter to calculate the ratio necessary to match the microphone impedance to the amplifier OSI that is realistically achievable. Since few people are intense enough about the whole affair to bother measuring individual microphones, the convention that 200 Ω is a good midpoint for source impedance serves well. Variations between actual microphones make trivial differences in the larger scheme of things. The assumption that most bipolar input amplifiers have an OSI of between 5 kΩ and 15 kΩ indicates that the transformer ratio should lie somewhere between 1:5 and 1:8.7.

Many consoles use higher ratios (typically 1:10), probably in the naive belief that the noise advantage of a step-up input transformer stems from the free gain it affords. Although on a basic level it would seem to make sense that the less electronic gain needed the quieter the system must be, this fallacy is completely belied by the truth that the transformer merely allows you to choose and alter the source impedance for which the amplifier is optimally quiet. Increasing the turns ratio beyond this easily defined optimum can and will actually render the amplifier noisier.

In practice the free gain can be more of a nuisance than a benefit. It is not unusual for microphone inputs to receive transients exceeding +10 dBu and mean levels of −10 dBu, especially in a rock-and-roll stage or recording environment. Even dynamic capsules can deliver frightening levels that can pose head room problems in the mixer front end. A typical 1:5 transformer has a voltage gain of 14 dB (20 dB for a 1:10 ratio), which would mean that even with no electronic gain after the transformer, normal mixer operating levels are being approached and possibly exceeded. These circumstances make worrying about a dB or two of noise performance total nonsense to be sure; it just serves to point out that our microphone front end has to be capable, if not perfectly optimized, for elephant herds as well as butterflies.

25.10.6.1 Transformer Characteristics

Transformers have numerous limitations and inadequacies resulting from their physical construction that make their actual performance differ (in some respects radically) from that expected of a theoretical model.

The heart of the transformer is the magnetically pliable material into and out of which energy is induced. Virtually any material—nickel, steel, iron, ferrous derivatives, and substitutes—have the same basic limitations. They saturate at a magnetic level beyond which they are incapable of supporting further excursion, and exhibit hysteresis—a crossover like nonlinearity at low levels responsible for a significantly higher distortion at low levels than anything else likely to be found within a well-designed modern-day signal path.

These two effects at opposite ends of the dynamic spectrum mean that all transformers have a well-defined range within which they must be operated and this range is less than the range of levels the microphone amplifier (mic-amp) is expected to pass. This is especially true at low frequencies, where the core is prone to saturation far earlier. Optimization begins here. Is it to be designed for minimum hysteresis (butterflies) or with plenty of material to be tolerant of monstrous (elephantine) signal levels?

Windings are made of wire, which has resistance. Resistance means loss and decreased efficiency and noise performance. By the time there are enough turns on each of the windings to ensure the inductive reactances are high enough not to affect in-band use, winding resistances can no longer be ignored.

Capacitance exists between things in close proximity and that includes transformer windings—between each other, between adjacent turns and piles in the same winding, and from the windings to ground. In this given instance it is nothing but bad news. Capacitance between windings means unwanted leakage and imperfect isolation, while winding self-capacitance reacts with the winding inductances to form resonances. Resonances, even if far outside the audio band, invite response trouble, and disturb in-band phase linearity. Combinations of these capacitances greatly affect one of the greater advantages of transformers, common-mode rejection (CMR).

25.10.6.2 Common-Mode Rejection

For a transformer to work and transfer wanted information from one winding to another, a current must be made to flow through the primary; this is ideally achieved just by the opposing polarity (differential) signal voltage applied across it. Again ideally, any identical signals on the two ends of the windings (common mode) should not cause any current to flow in the winding (because there is no potential difference across the winding to drive it) and so no signal transfer can be made into the secondary. So much for ideal.

Common-mode rejection is the ability of the transformer to ignore identical signals (in amplitude and phase) on the two input legs and not transfer them across the secondary as differential output information.

Principally, it is imbalanced distribution of capacitance along the length of the two windings, both with respect to each other and to ground that makes CMR less than perfect. Co-winding capacitance has the effect of directly coupling the two wiring masses permitting common-to-differential signal passage, which worsens with increasing frequency at 6 dB/octave. Electrostatic shielding (a Faraday shield) between the windings can alleviate co-winding capacitance coupling.

Further CMR worsening can be expected even if the two windings are perfectly balanced with respect to each other, if the primary winding is not end-to-end capacitatively matched with respect to ground. Any common-mode signal from a finite impedance source (almost always the case) when confronted with such a capacitatively unbalanced winding sees it as being just that—unbalanced (becoming more so with increasing frequency). Again, input common-mode signals are transferred across to become output differential information indistinguishable from the wanted input differential source.

Broadcasters particularly are concerned with winding balance, not only on microphone transformers but also on line-output transformers, reasoning that common-differential transference is as likely to occur at a source as at an input.

25.10.6.3 2Microphone Transformer Model

Fig. 25-40 gives a better idea of what the small signal of a dynamic microphone has to suffer. The winding capacitances (C_P and C_S) form lovely resonances with the inductances, while the transformed up primary winding resistance (R_P) added to the resistance of the secondary winding (R_S) merely serves to increase the effective source impedance of the microphone producing loss and resultant inefficiency.

A frequency response of a less-than-ideal transformer fed from a 200 Ω source and measured at high impedance across the secondary looks something like Fig. 25-41, where the low-frequency droop is attributable to one or both of the winding inductive reactances becoming comparable to signal impedances, while the high-frequency peak is an aforementioned secondary winding self-resonance. Usually the primary self-resonance is fairly well damped by the source impedance, but occasionally added cable capacitance can play cruel tricks here, too.

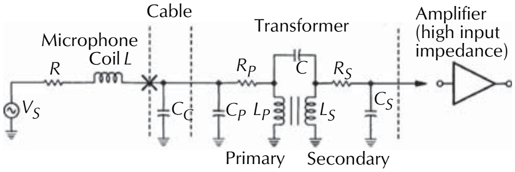


Figure 25-40. Transformer coupling model showing major elements.

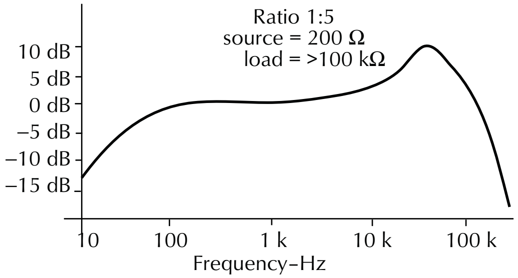


Figure 25-41. Typical transformer transmission response.

The mic-amp itself, as discussed, has a high-input impedance (hundreds of kilohms and up) while its optimum source impedance is defined at around 5–15 kΩ.

It's good engineering practice to consider how the circuit behaves when the operating impedances are no longer defined by the microphone (i.e., when it is unplugged). Ordinarily, the circuit of Fig. 25-40 with the microphone disconnected would probably oscillate, as would any circuit with a high-gain, high input-impedance amp terminated only by the collection of vile resonances and phase-shifting elements that are an open-circuit transformer. An open-circuit impedance-defining resistor (R_o in Fig. 25-42) with a value 10 or 20 times that of the amp OSI, helps tame this. It also marginally tames the secondary resonance.

There are a variety of techniques for dealing with this resonance. They vary from pretending it doesn't exist to actually using it as part of a front-end, low-pass filter to keep ultrasonic garbage out of the electronics. Minimization of the high-frequency bump is attempted as much as possible passively, prior to the amp; the taming network in Fig. 25-42 represents a typical approach. Here, a series resistor-capacitor combination in conjunction with the open-circuit impedance-defining resistor is used. The values are calculated to produce a step-type response, Fig. 25-43, which when combined with the hump at the high-frequency end of the transformer response, produces a more acceptable roll-off characteristic. Naturally, the interreaction between this network and the complex impedance of the transformer is not quite that simple. The network capacitance reacts heavily with the transformer inductance, shifting the resonance frequency in the process. It is this fact that has led to the misconception that the capacitance somehow magically tunes out the resonance.

Open-circuit stability is dramatically improved, Fig. 25-43. The network takes an even larger slice out of the overall high-frequency response, keeping impedances at the top end comfortably low.

25.10.6.4 Bandwidth

Providing the compensating high-frequency roll-off around a subsequent amplifier, in the form of exaggerated feedback phase-leading around the mic-amp itself in this case (C_F), has the advantage that the noise performance of the combination at higher frequencies remains unimpaired by an impedance mismatch resulting from a passive network.

Problems result in several areas. Compensation around the mic-amp becomes limited when the electronic gain approaches unity, while compensation around a late fixed-gain stage means that all stages prior to it, including the mic-amp, have head room stolen at the frequency of the resonance and to a degree of the magnitude of the resonance. This may or may not be a problem depending on how far the lower side of the resonant curve invades the audio band.

The passive method reduces the magnitude of the resonance. The ultimate low-pass roll-off slope is that of the high-frequency side of the resonance, which more accurately is a lightly damped inductance-capacitance, low-pass, 12 dB/octave filter. The active method uses an additional 6 dB/octave curve in the compensation making a total of 18 dB/octave, but it relies on the resonance being of a manageable degree to begin with. A measure of both techniques is usually required; their balance and relationship are an experimental process to optimize for each different type of transformer.

This enforced filtering is of considerable advantage, helping to keep all sorts of unwanted ultrasonic noise from finding its way into the mixer. It also represents a major advantage of transformer inputs over solid-state varieties.

A further advantageous filtering is the falling source impedance seen by the amplifier at extreme low frequencies. This is due to the fact that the winding inductive reactance reduces with frequency. This is a definite help in combating the generation of excess low-frequency noise in the first amplifier.

25.10.6.5 Common-Mode Transfer

There are two different amplitude response curves to be considered. The first, the normal differential input, has been fairly thoroughly determined. The second, by virtue of its mechanism, relies on imperfections within the main filter element itself (the transformer) rides over and oblivious to our carefully calculated filter

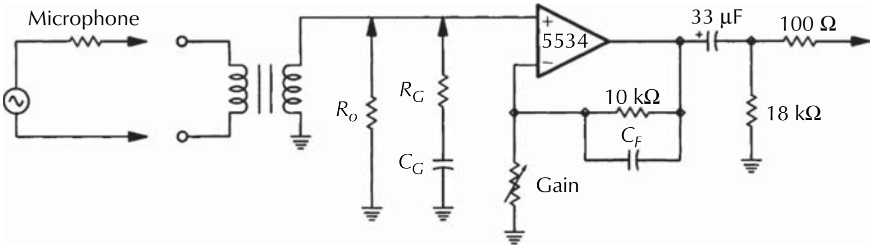


Figure 25-42. Basic microphone preamplifier with compensation components.

responses. Common-mode unrejected signals still appear at the amplifier input as if nothing had happened.

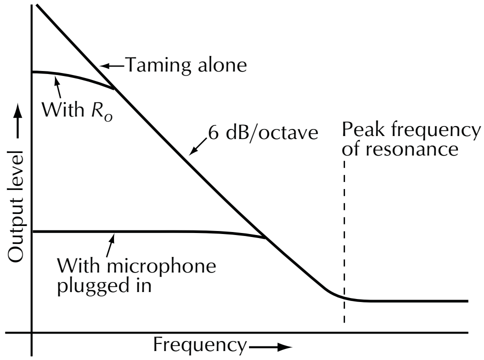


Figure 25-43. Frequency response of taming network.

25.10.6.6 Input Impedance

As determined earlier, we would end up with better noise performance and cleaner sounds if the microphone looked into a high, preferably infinite, impedance. Preferences apart, we have already had to define the reflected load (input), impedance by the resistor needed to keep the front-end stable under unplugged conditions (R_0), but at least it is an order of magnitude or so above working impedances, so its effect is small. It does, though, act as part of an attenuator of input signals along with the source impedance and winding losses, Fig. 25-44. This is the major factor responsible for worsening front-end noise performance using transformers. Any attenuation before the optimized amp directly degrades the noise figure, typically between 1 and 6 dB, depending on the transformer.

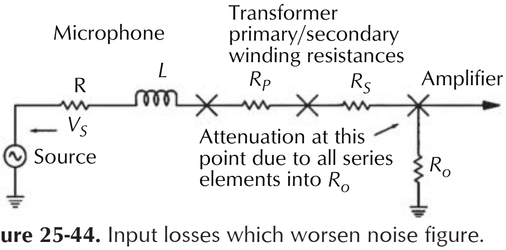

Microphone
Transformer primary/secondary winding resistances
R
L
R_P
R_S
Amplifier
V_S
Source
Attenuation at this point due to all series elements into R_o
R_o

Figure 25-44. Input losses which worsen noise figure.

If the transformer were perfect, it could be assumed that the reflected impedance, as seen by the microphone, would be constant over the audio band. At the low-frequency end, Fig. 25-45, the diminishing inductive reactance of the transformer windings (tending to zero with frequency) becomes a term of greater importance, affecting parallel impedances, attenuation, and accordingly, efficiency. Winding self-capacitances and the passive compensation networks are largely to blame for the high-frequency droop, although the list of contributing mechanisms is nearly endless.

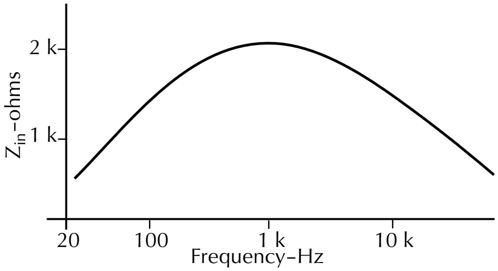


Figure 25-45. Typical input impedance curve.

A good rule of thumb is that the midband input impedance should exceed ten times the source impedance, or about 2 kΩ for a dynamic microphone. Any wild variation in this impedance is obviously going to result in frequency and phase response aberrations, which are probably the greatest single drawback to transformer front ends. Things aren't quite as bad as they seem; examples of performance shown here have been deliberately of a marginal transformer to highlight the illeffects, notably in response and impedance flatness; good transformers from good reputable manufacturers such as Jensen, Lundahl, and Sowter generally show much nicer results, but the design criteria to eke the best from them remain nonetheless.

25.10.6.7 Attenuator Pads

Attenuator pads, regrettably necessary in many instances to preserve head-room and prevent core saturation with elephantine sources, should maintain expected operating impedances when introduced. The transformer primary should still be terminated with a nominal 200 Ω, while the microphone should still look at 2 kΩ or above. Departure from these will cause the microphone/amplifier combination to sound quite different when the pad is thrown in and out, as would be expected from altering source and load impedances in and around complex filter characteristics. A significant downside to pads is that although the differential (desired) signal is being attenuated to the expected degree, any common-mode signal isn't.

25.10.6.8 Transformerless Front Ends

Bringing the amplifier optimum source impedance down to that of conventional dynamic microphones is possible by means other than transformers. Reducing the ratio of amplifier-inherent voltage and current noises has this effect. Two main techniques, either alone or in concert, are used:

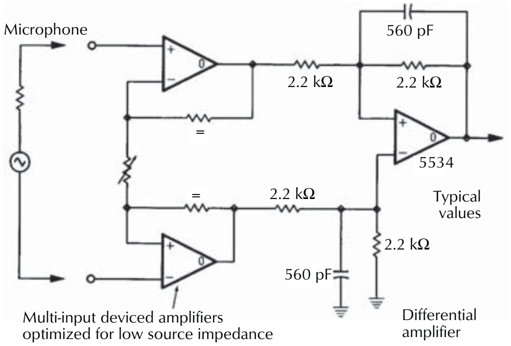
1. Large-geometry devices have innately lower noise impedances. Even power-amplifier drivers have been used (e.g., 2N4918, BD538) but these tend to suffer from low transit frequencies (bandwidth) and beta (gain), which can lead to additional complexity in the circuitry.
2. Paralleling multiple identical input devices, and so proportionally increasing the noise current in relation to the noise voltage, reduces the ratio between them (i.e., noise impedance).

PNP transistors, as mentioned elsewhere, have less surface-recombinant noise/lower base-spreading resistances than NPNs and are favored in this application.

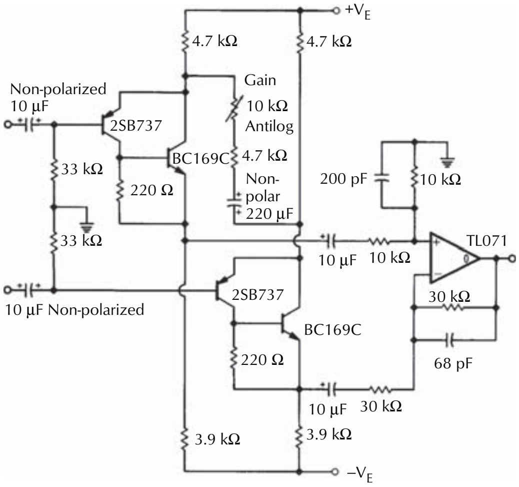
The usual technique is to place two of these large and/or multidevice input front-end amps—preferably accurately matched—ahead of an electronic differential amplifier, as shown in Fig 25-46. All the amplifier gain is made within the first pair of stages, differentially cross-coupled. This gain arrangement, rather than referring to ground, can afford reasonable common-mode signal rejection. Differential input signals are amplified since the reference for each of the two amplifiers is the other amplifier, tied to an identical signal of opposite polarity.

If the input signals to the two amps are identical in phase and amplitude (common), the references for each of the amplifiers are similarly waving up and down sympathetically to the signal. There is no voltage difference for the individual amplifier to amplify; consequently, there is no gain. For ordinary differential input signals, the amplifiers operate conventionally, their ground reference being a zero voltage point half-way along the gain-determining variable resistor. This point is a cancellation null between the opposite sense polarity swings of the two amplifiers.

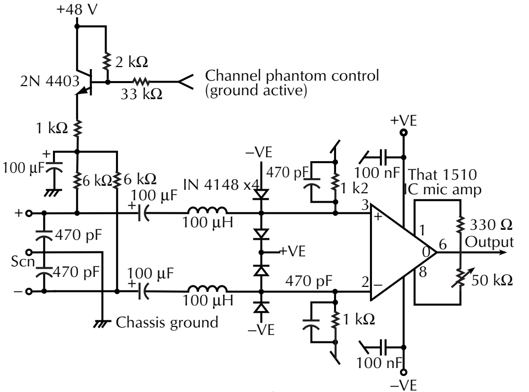
These amplifiers feed a conventional electronic differential amplifier running usually at unity gain. In order to maintain stage noise as low as possible, the resistors are made as low in value as the devices can sensibly stand. This arrangement is unmistakably a bastardized instrumentation amplifier—a well-documented circuit configuration; the only thing of remark is the pair of low-impedance optimized front-end stages.


Microphone
560 pF
2.2 kΩ
2.2 kΩ
=
5534
=
2.2 kΩ
Typical values
2.2 kΩ
560 pF
Multi-input deviced amplifiers optimized for low source impedance
Differential amplifier

A. Microphone-amplifier arrangement.


+V_E
4.7 kΩ
4.7 kΩ
Gain
Non-polarized 10 µF
10 kΩ Antilog
2SB737
BC169C
4.7 kΩ
33 kΩ
Non-polar 220 µF
200 pF
10 kΩ
220 Ω
33 kΩ
TL071
10 µF
10 kΩ
10 µF Non-polarized
2SB737
30 kΩ
BC169C
68 pF
220 Ω
10 µF
30 kΩ
3.9 kΩ
3.9 kΩ
–V_E

B. Simple discrete input.


+48 V
2 kΩ
2N 4403
33 kΩ
Channel phantom control (ground active)
+VE
1 kΩ
–VE
100 µF
6 kΩ
6 kΩ
IN 4148 x4
470 pF
1 k2
100 nF
That 1510 IC mic amp
100 µF
3
+
100 µH
1
330 Ω
470 pF
6
Output
+VE
Scn
470 pF
100 µF
470 pF
2
8
50 kΩ
100 µH
1 kΩ
Chassis ground
–VE
100 nF
–VE

C. Integrated mic amp.

Figure 25-46. Basic transformerless microphone-amplifier arrangement.

A criticism (rightly) leveled at some implementations of this, including earlier-generation integrated parts, is that the noise performance is significantly worse at low amounts of gain than at high gains, where the (hopefully) optimized input pair reigns. This is mostly down to the impedances chosen around the op-amp differential amplifier; some resistor values here have been seen as high as 25 kΩ, guaranteeing a high invariant noise floor for much of the gain range; bizarrely, some implementations attempt gain around this stage, too. The lesson earlier learned of keeping the circuit impedances as low as circumstances permit is salient here.

There are few reasons now for hand-knitting discrete versions of this input arrangement, and good ones not to, primarily suitable input devices: the much-favored 2SB737 is sadly becoming difficult to procure, and the excellent integrated matched transistor sets such as the THAT 320 are relatively expensive. Relative, that is, to the cost of a purpose-designed IC! There are a few integrated versions of this kind of arrangement offering very acceptable performance in the convenience of a little package; the Burr-Brown INA103 or 163 and the THAT Corporation's 1510 have multiparallel input transistor stages presenting OSIs about perfect for nominal microphone impedance, the latter part taking to heart the op-amp differential stage noise issue, with excellent results. Texas Instruments/Burr-Brown's PGA2500 is a digitally gain-controlled version of this configuration, which works very well indeed with fine gain steps and inaudible gain-change (zipper) noise, solving a major headache for digital console designers and others desirous of remote control of mic gain. They have proved themselves worthy with ribbon microphones, the ultimate front-end-noise torture test.

With this configuration, although potentially offering far higher and flatter input impedances than transformer inputs, there are, as always, snags. Common-mode signals directly gobble up head room in the first pair of stages even if they are operating as followers; that this common-mode stuff is substantially canceled in the following differential amplifier is a bit of a stable-door and bolted-horse routine. There is also the great danger that common-mode signals (in addition to normal differential signals) can exceed the input swing capability of the input devices. At best this will block the input stage, at worst—if the common-mode signal is big enough and at a low enough impedance (think ac line grounding fault here)—serious destruction can result. A transformer input would blithely ignore such.

Radio frequencies simply adore base-emitter junctions, and this configuration has them in abundance. Successfully filtering microphone inputs sufficiently without sacrificing noise performance or input device high-frequency gain (increasing high-frequency distortion and so on) is not a trivial task; it makes the self-filtering properties of an input transformer seem rather appealing.

Fig. 25-46C details the kind of helmet-and-armor with which one has to attire electronic microphone front ends to survive the fray:

- The π□L/C filters and additional coupling to ground at high frequencies help against RF. The inductors can either be single pieces, one in each leg, or preferably a pair of windings around a common (usually toroidal) core. This has the twofold advantages that the choking effect is concentrated on common-mode signals—the most common (so to speak) interference method—and that the inductances of the two windings essentially cancel for differential signals, so that there is much less effect of the RF protection impinging on the desired audio.
- The input dc-decoupling capacitors have to be pretty huge in value to maintain low-frequency integrity and at the same time have a high enough voltage rating to handle typically 48 V off-load phantom-power voltage.
- The parametric varactor capacitance of the clamping diodes has little to no effect on the audio but are vital to protect the device front end from the very healthy whack as the phantom voltage is turned on or off. Even so, clamped or no, good common-mode rejection or no, phantom ramped or no, these transitions are not exactly subtle. These diodes may help protect the input from other nasties, too, but the aforementioned major ground fault would render all this toast anyway. The good news: such are rare but in the real world of touring sound not unthinkable.

25.10.6.9 Line Level Inputs

The reader is referred to Chapter 11 for Bill Whitlock's excellent further coverage of real-world interfacing.

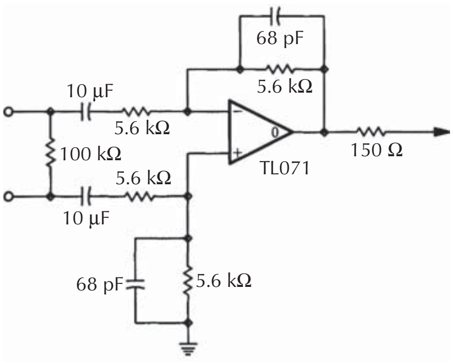
High-level balanced interconnections and systems for the most part have been largely relegated to outside world and intersystem interfacing; internal interconnects are left unbalanced except within a very few completely balanced console designs. The wisdom until quite recently was that balancing implied transformers and their performance limitations.

Transformers—good ones at least—are expensive, large, and heavy. Not so good ones are still relatively expensive, large, and heavy and represent a weak link in a modern signal path, with their low frequency distortion, hysteresis, and high/low frequency/phase response effects. Transformers are best used only where their impedance transformation capabilities, innate filtering, and excellent isolation properties are really needed. And then only really good ones. Most high-level interfacing applications within the confines of a studio environment justify neither their capabilities nor drawbacks.

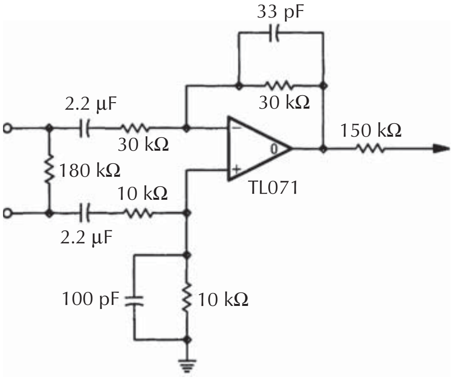
The search was on for electronic equivalents to transformers for both input and output applications, a moderate degree of success being achieved early on for input stages with classic circuits such as Figs. 25-47 and 25-48. These are simple differential input amplifier and instrument amplifier configurations using op-amps.

Line inputs are commonly differential amplifiers, similar to the one used in the transformerless mic-amp, but with the resistor values elevated to bring the differential input impedance up to over the 10 kΩ required of a bridging termination, Fig. 25-47. The noise of these stages is directly attributable to these resistor values, so the lower resistor values are better. An instrumentation amplifier configuration would seem to offer possibly better performance for noise (the differential amplifier resistor values may be kept small) but it entails the use of undesirable voltage followers (see Section 25.7) with potential stability problems, input voltage swing limitations, and unprotected (for RF) input stages. At least with a simple differential amplifier the impedances are comfortably low and the inputs buffered by resistors from the outside world.

The dc-blocking series capacitors must, unfortunately, be large in value to maintain an even input impedance and sensibly flat phase response at the lowest used frequencies. Also, being necessarily unpolarized, they are physically large and expensive. This is a small price to pay, though, for such a simple but important circuit element.


68 pF
5.6 kΩ
10 µF
5.6 kΩ
100 kΩ
5.6 kΩ
10 µF
0
TL071
150 Ω
68 pF
5.6 kΩ

A. Classic, equal value.


33 pF
30 kΩ
2.2 µF
30 kΩ
180 kΩ
10 kΩ
2.2 µF
0
TL071
150 kΩ
100 pF
10 kΩ

B. Optimized for matched dynamic amplifier.

Figure 25-47. Electronic differential input amplifier.

The instrumentation amplifier presents very high, nonground-referred differential and common-mode terminations and has the great advantage that gain may be easily invoked between the two input amps at no cost to the excellent common-mode rejection, Fig. 25-48. Integrated line receivers are typically of this configuration.

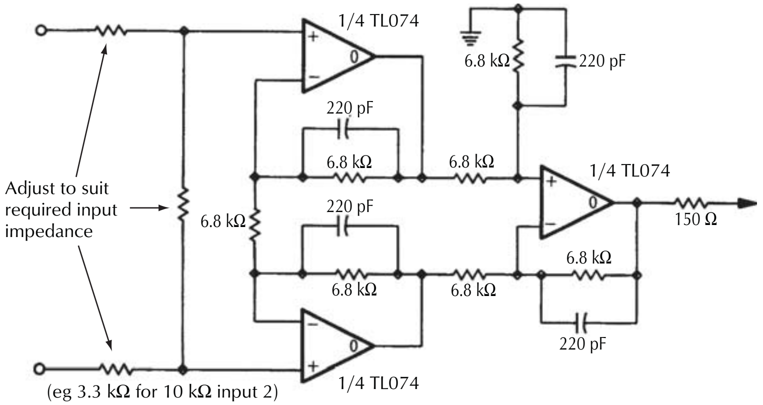

1/4 TL074
0
6.8 kΩ
220 pF
220 pF
6.8 kΩ
6.8 kΩ
1/4 TL074
0
Adjust to suit required input impedance
6.8 kΩ
220 pF
150 Ω
6.8 kΩ
6.8 kΩ
6.8 kΩ
220 pF
0
1/4 TL074
(eg 3.3 kΩ for 10 kΩ input 2)

Figure 25-48. Instrumentation amplifier-type line-input stage.

A pair of inverting amplifiers, shown in Fig. 25-49, provides a simple, hardy, easily defined differential (but not true floating balanced) input stage. A fascinating circuit known as the *Superbal* input is depicted in Fig. -50; this is a balanced differential virtual-earth amplifier, referred to ground solely by one op-amp input and capable of very good common-mode rejection, limited by the tolerance of the components from which it is constructed. Accepting any lopsided input signal, it delivers a differential output perfectly symmetrical to ground, making it an exceptionally useful input conditioning amplifier.

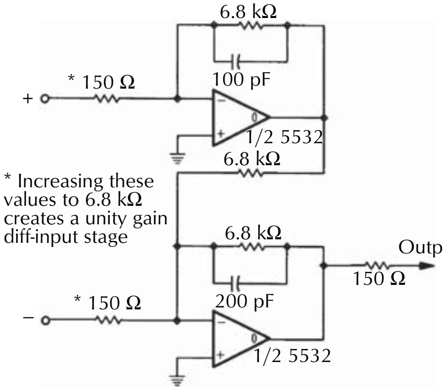


Figure 25-49. Differential mix/input amplifier.

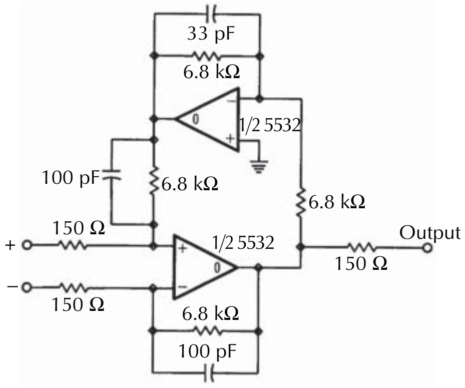


Figure 25-50. Superbal differential mix/input amplifier.

The capacity of both these circuits to be differential virtual-earth points makes them ideal for use in balanced mixing bus systems.

25.10.6.10 Electronic Balanced Outputs

The simplest balanced outputs configuration is given in Fig. 25-51. This is a pure, no-nonsense, inverter-derived differential feed. For many internal interconnections and especially in differential balanced mixing systems it works well, but it should not be used to connect to the outside world.

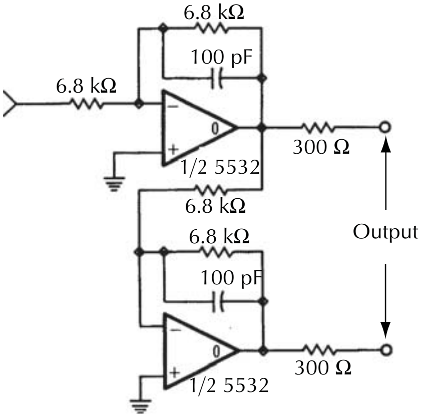
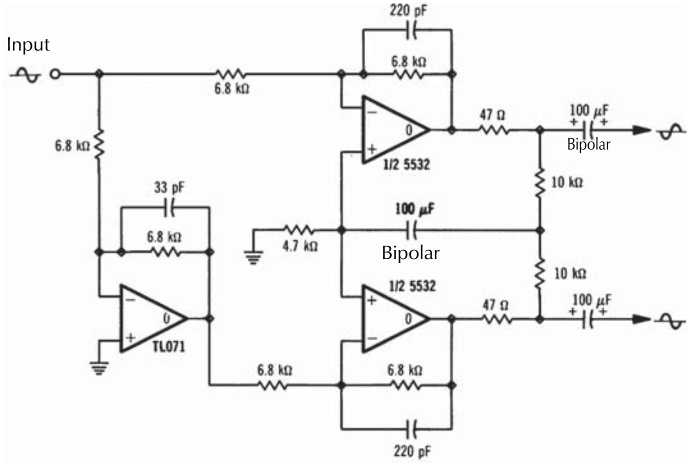


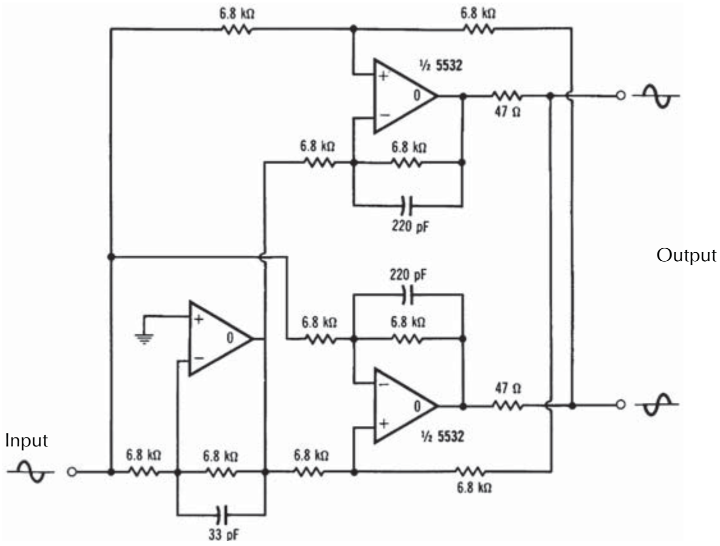
Figure 25-51. Inverter-type differential output.

Ideally, there must be no discernible difference in characteristics between the output circuit and an ideal transformer. After all, the fate of signals in the real world on a balanced transmission line won't alter in your favor simply because you've chosen not to use a transformer. If transformers are being supplanted it had better be with devices capable of affording similar benefits to the system and its signals. Regardless of applied reverse common- mode potential, the differential output potential must not change. Also the output should be insensitive to any imbalance in termination, even to the extent of shorting one of the legs to ground. This is the floating test. For example, the simple inverter circuit of Fig. 25-51 fails the floating test since, if one leg is shorted to common, the overall output has to drop by one-half (6 dB). (The question of what happens to ground noise with a shorted amplifier bucketing current into it will be sidestepped here.) Two basic circuits have emerged as being close approximations to a transformer. Not only are they fairly closely related, but most balanced output topologies are also derived from them. They both depend on cross-coupled positive feedback between the two legs to compensate for termination imbalance.

In Fig. 25-52 a unity-gain inverting stage provides out-of-phase drive for the two legs, each output leg of which is a −6 dB gain inverting amplifier with error sensing applied to its reference (positive) inputs. Under normal operation, there is no error-sensing voltage; the two inverse outputs cancel at the midpoint of the equal sense resistors. The two amps invert a differential


Input
220 pF
6.8 kΩ
6.8 kΩ
6.8 kΩ
47 Ω
100 µF
Bipolar
1/2 5532
10 kΩ
33 pF
6.8 kΩ
100 µF
4.7 kΩ
Bipolar
1/2 5532
10 kΩ
47 Ω
100 µF
TL071
6.8 kΩ
6.8 kΩ
220 pF

A. Inverter type.


6.8 kΩ
6.8 kΩ
½ 5532
47 Ω
6.8 kΩ
6.8 kΩ
220 pF
Output
220 pF
6.8 kΩ
6.8 kΩ
47 Ω
½ 5532
Input
6.8 kΩ
6.8 kΩ
6.8 kΩ
6.8 kΩ
33 pF

B. Cross-coupled differential amplifier type.

Figure 25-52. Electronic floating differential amplifier stage.

voltage equal to the unbalanced input voltage appearing between their outputs. (Two –6 dB quantities sum to make zero gain.) Take the case of one output, the upper one being shorted to ground. An error potential is derived of such a phase and level on the error-sense line that positive feedback increases the gain of the unshorted amp by 6 dB, while matching on the positive input of the shorted one the signal on the negative input, canceling its amplification. Closing the shorted amp down prevents ground-current problems; therefore, any measure of output termination imbalance is reasonably dealt with by this arrangement.

A major problem with any circuit depending on high levels of positive feedback such as these is their potential instability. Both these circuits are right on the edge of instability—they have to be in order to work accurately; a measure of margin has to be given for peace of mind and component tolerances. This backing-off compromise affects primarily common-mode rejection and output level against lopsided terminations. A loss of about 0.5 dB in differential output level can be expected when one side is shorted to ground, although tight component tolerances can improve on this. Component tolerance imbalance—even if constructed with 1%

resistors—manifests itself also as sometimes quite substantial dc offsets that will likely have to be trimmed to not eat up too much head room.

Curiously, instability tends to show itself as common mode. This fault manifested itself to the author for the first time (too) early one installation morning; a peak program meter (PPM) across such an output read nothing, listening elicited a little bit of hum, but a scope on either leg to ground showed 10 Vp-p square waves, driving the tape machine to which the output was connected into shock.

Integrated versions of the cross-coupled circuit such as the SSM 2142 have the great advantage of extremely closely matched/trimmed resistor values, and hence far more predictable performance than discrete versions.

25.10.7 A Practical Microphone-Amplifier Design

Optimizing front-end sound is nothing more than shrewd judgment in juggling the nearly endless electronic operating conditions so that adequate performance is obtained over the wide range of expected and common input signals. Any wrinkles should be arranged to exert influence only under quite extraordinary operational conditions.

The microphone amplifier example described here, Fig. 25-53, is a somewhat developed version of a basic front-end design and is in grave danger of becoming an industry standard.

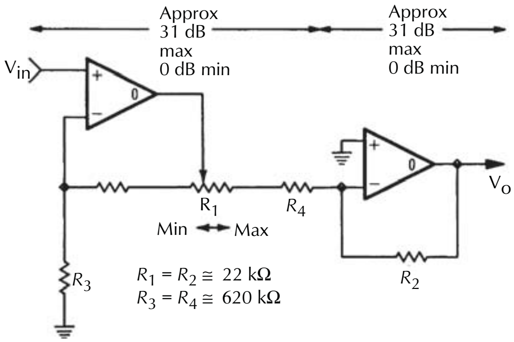

Approx 31 dB max 0 dB min
Approx 31 dB max 0 dB min
V_{in}
+
−
0
+
−
0
V_O
R_1
R_4
Min ←→ Max
R_3
$R_1 = R_2 \cong 22\ k\Omega$
$R_3 = R_4 \cong 620\ k\Omega$
R_2

Figure 25-53. Shared-gain two operational-amplifier input stage.

Initially most striking is the manner in which a single-track potentiometer is used to vary simultaneously the gains of two amplifying elements—the front-end (noninverting) stage and the succeeding inverting amplifier. Since the first stage is (as far as its inputs are concerned) a conventional noninverting amplifier, transformer input coupling is no more problematic than with simpler microphone amplifiers (e.g., Fig. 25-41, a standard generic microphone amplifier).

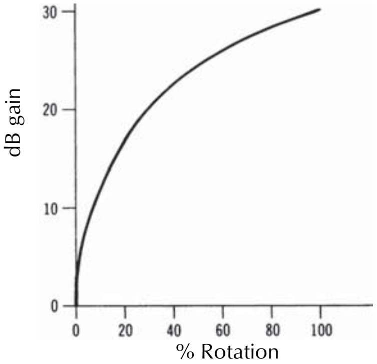
With maximum gain distributed between two stages, large gain is possible without any danger of running out of adequate steam at high frequencies for feedback purposes in either of the two amplifiers. This, incidentally, also makes for reasonably simple stabilization of the amplifiers, something not easily accomplished with simpler single-amplifier circuits achieving the same gain swing. Other than the obvious simplicity and economy of one-pot gain control, two nice features inherent in the design are interesting from the points of view of system-level architecture and operation.

25.10.7.1 System-Level Architecture

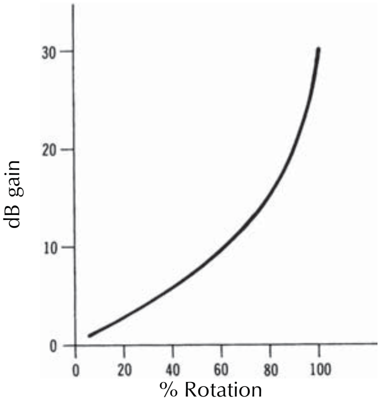
System-level architecture is largely concerned with operating all the elements of a system at the optimum levels and/or gain for noise and head room (i.e., at a comfortable place somewhere between the noise floor and clipping ceiling). Where gain is involved, it's important that the resultant noise be due primarily to the gain stage that has been optimized for noise (or rather lack of it) such that it can then mask all subsequent and hopefully minor contributions. At no point in the gain swing—particularly at minimum gain—should it be necessary to attenuate unwanted residual gain. This amount of attenuation gets directly subtracted from overall system head room. What good is 24 dB of head room everywhere else, if you have only 16 dB in the front end?

In this respect circuits similar to Fig. 25-53 score well, and the graphs of Fig. 25-54 show why. Fig. 25-54A represents the gain in dB of a simple noninverting amp varying with the percentage rotation of an appropriately valued linear pot in its feedback leg. This is like the gain/rotation characteristic of the first amp of Fig. 25-53. Similarly, Fig. 25-54B is the gain/rotation plot for a linear pot as the series element in an inverting amp, such as the second gain stage of Fig. 25-53. For the first half of the rotation, the first stage provides all the gain swing and most of the gain; only about 6 dB is attributable to the inverting stage at midpoint. Toward the end of the rotation, this position reverses with the front end remaining comparatively static in gain; the extra swing and gain come from the inverting stage. Noise criteria are met, since the first (optimized) stage always has more than enough gain to allow its noise to swamp the second stage, with the exception of minimum gain setting. There it hardly matters anyway because the front-end noise contribution is going to be at a similar level to the overall system noise floor (i.e., really quiet!). The impedances around the second stage largely determine the noise performance of the amplifier, and this is such that it need not be considered in relation to input noise at any sensible gain setting. Head room is satisfactory because no attenuation after the first gain stage is needed for any gain setting. The two gain stages operate nicely complimentarily.

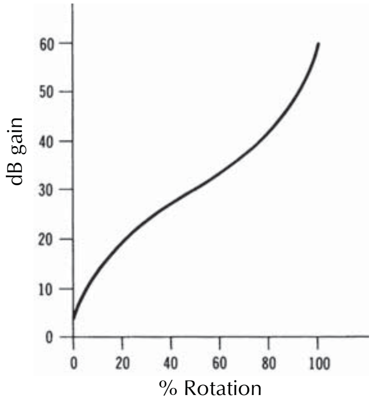
An operational advantage can be gleaned from Fig. 25-54C. This is the combined gain/rotation curve for the entire two op-amp circuit. Note that for a very large percentage of rotation around the middle of the gain swing (where it's most often used) the dB gain change per rotation is as good as linear. It gets a bit cramped at the top and bottom, but you can't win them all. For reference a little later on, it may be noted that there are two available resistors (R_2 and R_3) that may be used to modify the gain structure independently of the potentiometer.



A. Noninverting amplifier gain.



B. Inverting amplifier gain.



C. Combined noninverting and inverting gains.

Figure 25-54. Gain versus pot rotation for two op-amp input stage.

25.10.7.2 Input Coupling

As a microphone amplifier, the fairly high optimum source impedance of the op-amp used in Fig. 25-55 (a Signetics NE5534, or AD797) needs to be matched to the likely real source impedance of some 150–200 Ω. No apologies are offered for the use of transformer input coupling, as grossly unfashionable as this may currently seem. Transformers still offer outstanding advantages—especially simplicity, impedance step-up, protection, and filtering—over electronic inputs in this application.

Many circuit values (marked with an asterisk in Fig. 25-55, with some in quite unexpected places) are dependent on the specific transformer type in use. Several differing transformers can be very successfully used provided their differing ratios are taken into account in level calculations; a ratio of 1:7 is optimum to match the OSI of the input device employed. Phase and response trimming values will vary significantly. For example, with the Jensen JE-115-K, it is simpler than with the Sowter 3195 around which this circuit was originally developed. Despite the apparent simplicity of the circuit, a lot of effort has gone into defining the front-end bandwidth and straightening out the phase response at audible extremities. Taming the high-frequency transformer resonance in particular is quite tiresome.

On the front of the transformer hang the usual components to make the microphone amplifier useful in this world of capacitor microphones: a 20 dB input attenuator and 48 V phantom power via matched 6.8 kΩ resistors per leg carried common mode along the microphone line. Further to earlier discussions, the component values in the pad are chosen such that the microphone still sees the same general impedance whether or not the pad is inserted, while the mic-amp still sees about a 200 Ω source to keep all the transformer-based filtering in trim. It is essential the 6.8 kΩ phantom resistors are

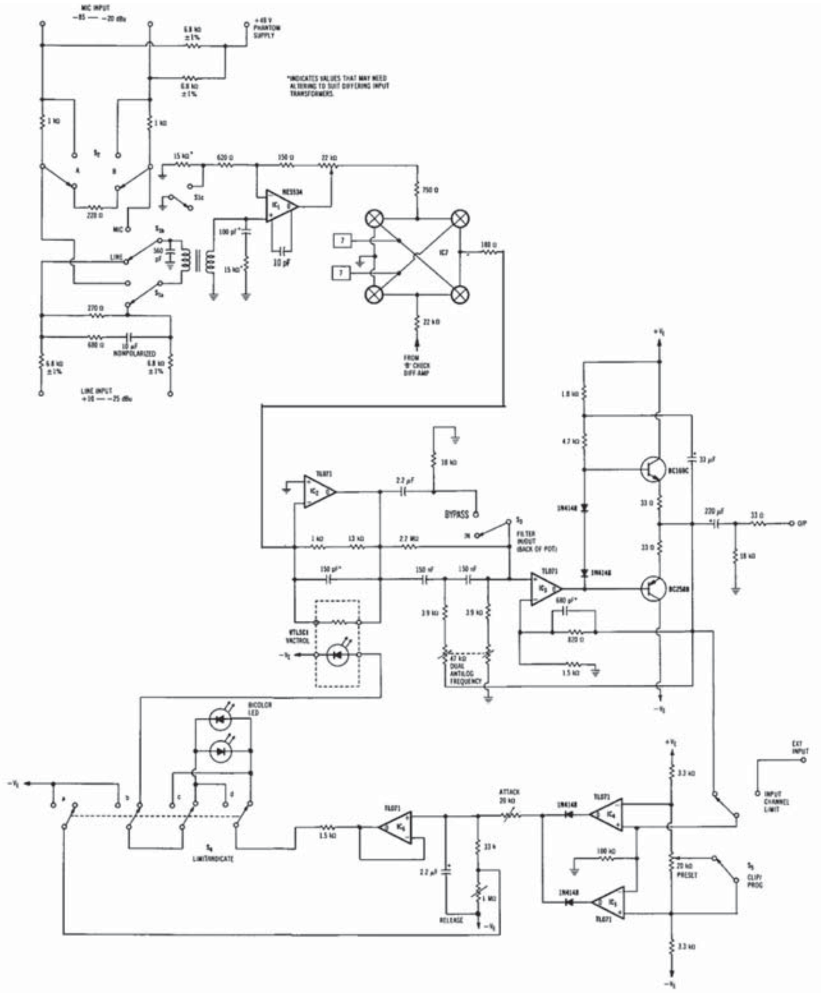

MIC INPUT
−85 — −20 dBu
6.8 kΩ ±1%
+48 V PHANTOM SUPPLY
6.8 kΩ ±1%
*INDICATES VALUES THAT MAY NEED ALTERING TO SUIT DIFFERING INPUT TRANSFORMERS.
1 kΩ
1 kΩ
A
B
15 kΩ*
620 Ω
150 Ω
22 kΩ
S1c
NE5534
IC1
750 Ω
220 Ω
MIC
S1b
LINE
560 pF
100 pF*
15 kΩ*
10 pF
IC7
180 Ω
7
7
S1a
270 Ω
680 Ω
10 μF NONPOLARIZED
22 kΩ
FROM 'B' CHECK DIFF AMP
6.8 kΩ ±1%
6.8 kΩ ±1%
LINE INPUT +10 — −25 dBu
+V1
1.8 kΩ
4.7 kΩ
18 kΩ
TL071
IC2
2.2 μF
33 μF
BC169C
1N4148
33 Ω
BYPASS
S3
220 μF
33 Ω
O/P
IN
FILTER IN/OUT (BACK OF POT)
1 kΩ
13 kΩ
2.7 MΩ
33 Ω
18 kΩ
150 pF*
150 nF
150 nF
TL071
IC3
1N4148
BC258B
680 pF*
3.9 kΩ
3.9 kΩ
820 Ω
VTL5C1 VACTROL
−V1
47 kΩ DUAL ANTILOG FREQUENCY
1.5 kΩ
−V1
BICOLOR LED
EXT INPUT
+V1
3.3 kΩ
−V1
a
b
c
d
ATTACK 20 kΩ
1N4148
TL071
IC4
INPUT CHANNEL LIMIT
TL071
IC6
1.5 kΩ
S4 LIMIT/INDICATE
33 k
100 kΩ
20 kΩ PRESET
S5 CLIP/PROG
2.2 μF
1 MΩ
1N4148
IC5
TL071
RELEASE
−V1
3.3 kΩ
−V1

Figure 25-55. Channel input amplifier, high-pass filter, and limiter.

matched; poor matching is a very easy way to wreck carefully won common-mode performance.

25.10.7.3 Line Level Input Facility

A line-in option is brought in via the transformer also. It features far stiffer input attenuation (about 36 dB) while simultaneously disabling much of the gain swing of the first amp. The resultant gain swing of 35 dB (between −25 dBu and +10 dBu input level) with a bridging-type input impedance of some 13 kΩ should accommodate most things that the microphone input or machine-return input differential amp can't or won't. A small equalization network is used in the attenuator to bolster the extreme low-frequency phase response.

An alternative and in many ways preferable line-in arrangement might be to use either a discrete or integrated instrumentation amp-style stage, switched into the second stage of the mic-amp. In Fig. 25-55 this would come in where "from B-check Diff-Amp" is marked. This would avoid the necessity of using the transformer with attenuators, which of course does no favors to the common-mode rejection ratio.

25.10.7.4 Common-Mode Rejection Ratio

Common-mode rejection ratio (CMRR) in the transformer is dependent mostly on the physical construction of its windings. The Sowter, in common with many other transformers, may be in need of compensation by deliberately reactively unbalancing the primary winding to match the inadvertent internal characteristics, Fig. 25-56. Jensen transformers are uncannily good in this respect—no tweaks usually being necessary. There are external circuit influences that can and will upset the maximum obtainable common-mode rejection. The accuracy of the phantom-power resistors is one; any input pad, regardless of accuracy, is another. Assuming any reactive (i.e., rising with frequency) common-mode response has been trimmed out, unequal phantom legs will enforce a lopsided flat common-mode response while true floating input pads instantly reduce the CMRR by nearly the amount of their attenuation. Why? They do this because they only attenuate the differential (wanted) signal and not the common-mode one. A halfway solution is to centrally ground reference the pad. Given all that, less than perfect common-made response shouldn't cause any ill manifestations in a typical recording environment with fairly short input leads. A high electromagnetic field of any sort, or an application with very long leads (or worse yet, a multicore), is far more likely to create problems with untrimmed inputs than with those properly balanced; vulnerability is greatly increased to all types of common-mode problems including noise on the phantom power-supply feed. Indeed, this is a common compounding of faults on a console that exhibits consistently noisy inputs.

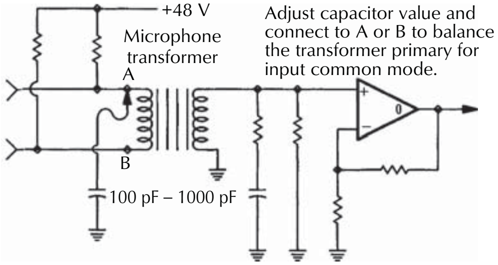

+48 V
Microphone transformer
A
B
100 pF – 1000 pF
Adjust capacitor value and connect to A or B to balance the transformer primary for input common mode.
+
−
0

Figure 25-56. Input common-mode "tweak."

25.10.7.5 Minimum-Gain Considerations

A minor compromise is necessary in the first stage to prevent its gasping with exhaustion on extremely high input levels. Ideally, the output of the operational amplifier has to look into an impedance of 600 Ω or greater (this being the lowest impedance into which it can drive full-output voltage swing). Maximum gain state isn't really a problem. If the first stage is overdriven, then the second stage will be some 30 dB into clipping; someone might notice!

At minimum gain though, the first stage in Fig. 25-55 is operating almost as a follower with an output load of 770 Ω, with the remaining feedback path to ground. That's safe and easily within the amplifier's driving capability. It would be better though if this small resistance were still smaller because it is contributing a little unwanted thermal noise to the otherwise beautifully optimized front end. The calculated degradation is only minor points of a dB and in practicality is easily lost in the gray mist that always surrounds the marriage of calculation with practical noise measurement.

The idea of using a front-end stage that turns into a follower under operating conditions has proved stable without any obvious trace of ringing within its bandwidth. This is probably because it is only being asked to look into safe, unreactive loads. Things that will make any unstable circuit squeal have not affected it. Among the instruments of torture have been a pulse generator/storage scope and an RF sweep generator/spectrum analyzer. The 10 pF compensation capacitor is more an act of conscience than a practical necessity. No compromise comes from its use here, since at maximum gain the first amp is working 30 dB below system level (an implied slew rate of nearly 200 V/µs!). At minimum gain the incoming signal level is such that it's most likely coming from a line source of certainly much more limited speed than the front end.

Down from the nether world of megahertz, the microphone amplifier is totally stable for audio, even with the microphone unplugged and input unterminated; the input network (of RG and CG) is designed to work in conjunction with the fairly low-input impedance of the 5534 (150 kΩ nominal).

25.10.7.6 The Limiter

Elaboration on the simple two op-amp mic-amp element consists of arranging an automatic gain control element in the feedback loop of the second amplifier and following that with a variable turnover frequency high-pass filter, Fig. 25-55.

A photoresistor device has its resistive end strapped across the normal gain-determining feedback resistor. Its resistance drops in value from very high (megohms) in inverse relation to the photodiode current to a limit of around 300 Ω at about 20 mA diode current. This resistance swing in the second amplifier is easily adequate for use in a peak limiter arrangement. The resistance change is close to exponential versus diode current, which could be of use in a gentler compressor, but here as a limiter the resistance change is quite sudden once that point is reached.

The limiter side chain is true symmetrical peak detecting, selectable to be able to pick off from either the high-pass filter output (as an input limiter) or from after the post-equalization breakpoint downstream (as a channel limiter). A positive-going and a negative-going level-detecting comparator are adjustable between clip detection (0.67 dB before system head room) or program level (nominally) +8 dBu.

A bicolor LED blinks red to indicate limiting in action, and it blinks green when the limiter is disabled to signify that the selected level (clip or program) is being reached or exceeded. In this indicate mode, the limiter integration time constant is deliberately shortened to make the green flashing similar in character to the red flashing in limit.

The difference is due to the nature of servo loops, of which a feedback limiter such as this is an example. In limit, the loop is self-regulating, the gain-control element holding back the audio level so that it's just tickling and topping up the side chain. In indicate, the loop is broken, and there is no such regulation. The green light stays on whenever the threshold is exceeded and tends to hang on while the time-constant capacitor discharges. With even a minor overdrive, this hangover could extend for quite a few seconds; hence, the shortened time constant.

This limiter is not subtle. The comparators deliver a full-sized, power-supply wallop to the integrator upon threshold, softened a bit by the attack preset in conjunction with the output impedance of the comparators. This rather unusual approach is to help wake up the photoresistor that has a relatively leisurely response time. The combination can be adjusted to be slow enough such that it doesn't clip yet fast enough to prevent an audible snap. Overshoot is generally within 1 dB on normal program, given a release time long enough to prevent pumping.

As a rough guide, if it's intended to use such a limiter for sporadic transient protection, it's best to aim for short attack and release times, bearing in mind that such settings will behave more as a clipper to the lower frequencies. For continual effect use, longer time constants will be less gritting and more buoyant. This side-chain arrangement certainly behaves differently from more conventional FET or voltage-controlled amplifier (VCA) linear proportional systems and needs a slightly different approach in setting up.

25.10.7.7 High-Pass Filters

Constructed around the line output amplifier of the front end in Fig. 25-55 is a second-order high-pass filter. It is a completely ordinary Sallen-Key type filter, arranged to use a dual-gang equal value potentiometer to sweep the 3 dB down turnover frequency from between 20 Hz and 250 Hz. A click-stop switch at the low-frequency end (counterclockwise) negates the filter, replacing it with a very large time-constant, single-order dc decoupler. These are both tied to reference in order to minimize clicks. Fortunately, the BiFET op-amp in the filter barely uses any input bias current, so there is little developed offset voltage from that source to worry about.

Being an equal-value filter, the *Q* or turnover would be very lazy indeed if the feedback were not elevated in level to compensate for the upset resistor ratio. Here a compromise is struck. A low *Q* gives a very gentle roll-off (which is sonically good), and high *Q* results in a much more rapid attenuation beyond the cutoff frequency at the expense of a more disturbed in-band frequency response—pronounced bumps—and frantic temporal and phase responses exhibited as ringing and smeared transients. Luckily, the majority of control-room monitors exhibit far worse characteristics at the low-frequency end.

A maximally flat response midway between the two extremes is chosen by an appropriate amount of elevated feedback (around 4 dB). This gain is taken across the filter as a whole, with the second stage of the microphone amplifier arranged to sustain a 4 dB loss to compensate. It all works out in the end, with no compromise of head room. With minimum gain set, there is still unity electronic gain front to back. An added convenience of gain is that it provides a better chance of shoring up feedback phase margin, which is quite important in a line amp that may have to drive a lot of heavily capacitative cable. Also, it provides yet another single-order low-pass pole to help smooth out the high-frequency resonance of the microphone transformer.

25.11 Equalizers and Equalization

The term *equalization* is strictly a misnomer. It was originally utilized to describe the flattening and general correction of the response of systems in which by a matter of course or design had deviated from the original shape (e.g., telephone lines and analog tape machines). (In the latter case, equalization refers to the adjustment tweaks to the preemphasis and deemphasis curves—not necessarily the curves themselves.)

In search of a name for the deliberate modification of amplitude and phase versus frequency responses for taste and the occasional genuine creative effect, the contraction EQ is well understood as both a noun and a verb.

This sonic mutilation uses frequency response curves and shapes in degrees that have grown through an uneasy mixture of operator needs and technical expedience/feasibility. One of today's multiparametric console channel EQs would have needed a rack full of tubes in the fifties and sixties. Funny, they didn't seem to need such EQs then.

The delight (and maybe curse) of IC op-amp design is that active filter (hence, EQ) implementation and techniques have blown wide open, limited only by economics, the largeness of the printed circuit board and the smallness of the user's fingers.

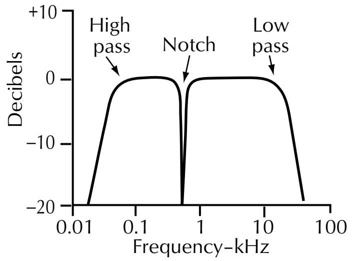
EQ curves can be roughly lumped into three user categories: garbage disposal, trend, and area. High-pass, low-pass, and notch filters that eliminate air-conditioning burble, mic-stand rumble, breath noises, hum, TV monitor line-frequency whistles, and excessive electronically generated noise are obviously in the business of garbage disposal. Fig. 25-57A shows the sorts of responses to be expected from these. Gentle hi-fi-type treble and bass slopes and similar shelving curves establish response trends shown in Fig. 25-57B, while resonance like, bell-shaped lift-and-cut filters manipulate given areas of the overall spectral response, Fig. 25-57D. These are used to depress unwanted or irritating aspects of a sound or, alternatively, to enhance something at or around a given frequency that would otherwise be lacking. As the curves differ, so do the design techniques required.

25.11.1 Single-Order Networks

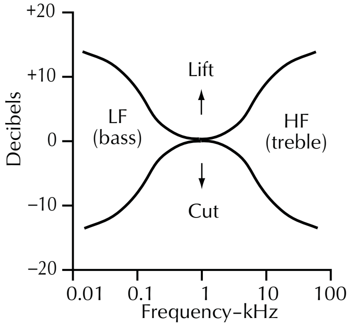
You can't build a house until you have the bricks, so they say. Fig. 25-58 has those bricks in the form of combinations of basic passive components with a rough guide to their input-output voltage transfer functions (essentially the frequency responses). Assumptions are that the V_{in} source impedance is zero and the V_o termination is infinite impedance.

Capacitative reactance decreases with increasing frequency, working against the resistance to increasingly short the output to ground with increasing frequency in Fig. 25-58A, while in Fig. 25-58B the capacitance steadily isolates the output from the input with reducing frequency (rising reactance).

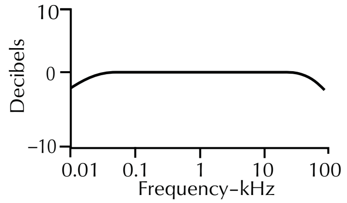
Inductors have entirely the opposite reactive characteristics. Inductive reactance is directly proportional to frequency, so the curves in Figs. 25-58C and D will be of no surprise at all, being complementary to those involving capacitance.



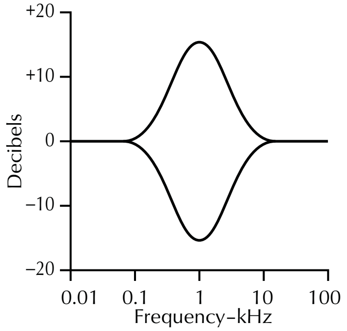
A. High-pass, low-pass, and notch typical responses.



B. Treble and bass shelving responses.



C. Flat response.



D. Bell or Gaussian responses.

Figure 25-57. EQ responses.

25.11.2 Single-Order Active Filters

Further useful curves are derived when the passive *R*, *C*, and *L* elements are wrapped around an op-amp in the classic inverting and noninverting amplifier modes, as shown in Figs. 25-58E to *L*. All the curves in Fig. 25-58 are normalized to unity gain and the same center

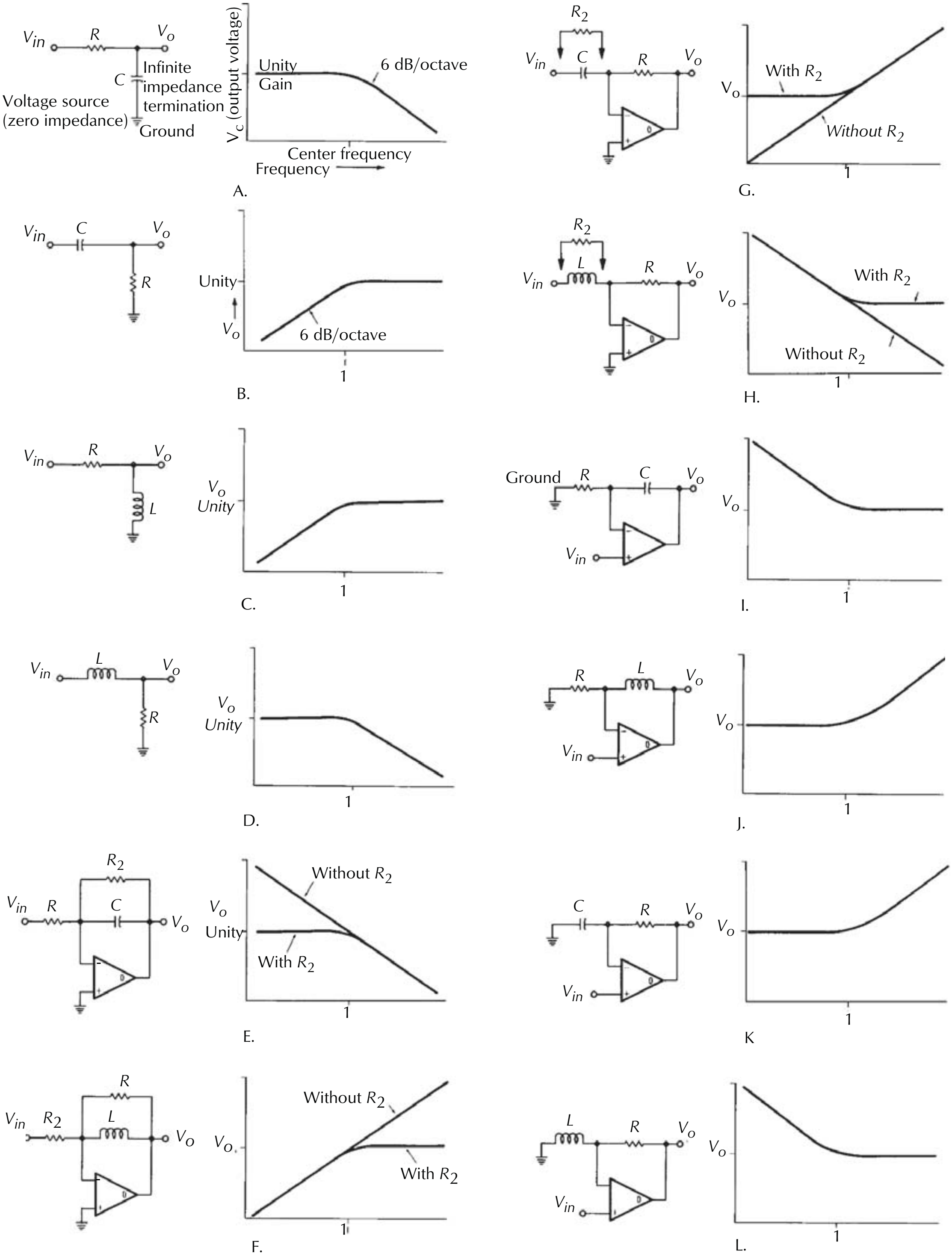

V_{in}
R
V_o
C
Infinite impedance termination
Voltage source (zero impedance)
Ground
V_C (output voltage)
Unity Gain
6 dB/octave
Center frequency
Frequency
A.
V_{in}
C
V_o
R
Unity
V_o
6 dB/octave
1
B.
V_{in}
R
V_o
L
V_o
Unity
1
C.
V_{in}
L
V_o
R
V_o
Unity
1
D.
R_2
V_{in}
R
C
V_o
V_o
Unity
Without R_2
With R_2
1
E.
R
V_{in}
R_2
L
V_o
V_o
Without R_2
With R_2
1
F.
R_2
V_{in}
C
R
V_o
V_o
With R_2
Without R_2
1
G.
R_2
L
V_{in}
R
V_o
V_o
With R_2
Without R_2
1
H.
Ground
R
C
V_o
V_{in}
V_o
1
I.
R
L
V_o
V_{in}
V_o
1
J.
C
R
V_o
V_{in}
V_o
1
K
L
R
V_o
V_{in}
V_o
1
L.

Figure 25-58. Single order filters.

frequency at which the curve departs significantly from flat.

Standard arithmetic formulas normally consider or obtain a frequency at which the curve has departed 3 dB from flat (the *3 dB down point*) being usually also where the phase has been shifted 45°. This is only partially useful in the design of filters for use in practical EQs; the departure point, or turnover frequency, is generally more relevant.

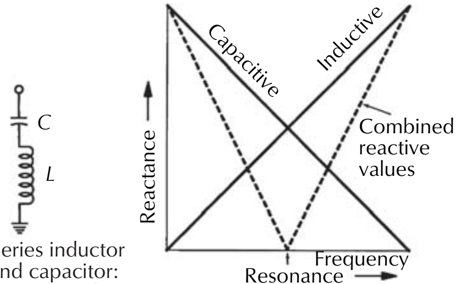
25.11.3 Changing Filter Frequency

With any of these filters, moving the frequency at which the filter bites can be achieved by altering any of the *R*, *L*, or *C* values. Making any value smaller moves the frequency higher, while making the value larger moves the frequency lower.

There are an endless number of combinations of element values to create the same curve at the same frequency. In Fig. 25-58A if the value of the capacitor were reduced (increased in reactance), the filter curve would shift up in frequency. A corresponding proportional increase in the series resistor value would result in the original turnover frequency being restored; we have an identical filter with a different resistor/reactor combination. What does remain the same is the ratio or relationship between the two elements. It is only the filter impedance (the combination of resistance and reactance) that varies.

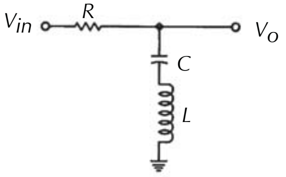
With the exception of a few, the operation of any active filter can eventually be explained by referring to these basic single-order filter characteristics in Fig. 25-58.

There is one particular combination of two reactive elements (capacitance and inductance) that is of prime relevance to the construction of EQs. This, a series-tuned circuit, Fig. 25-59, is where things really become interesting.

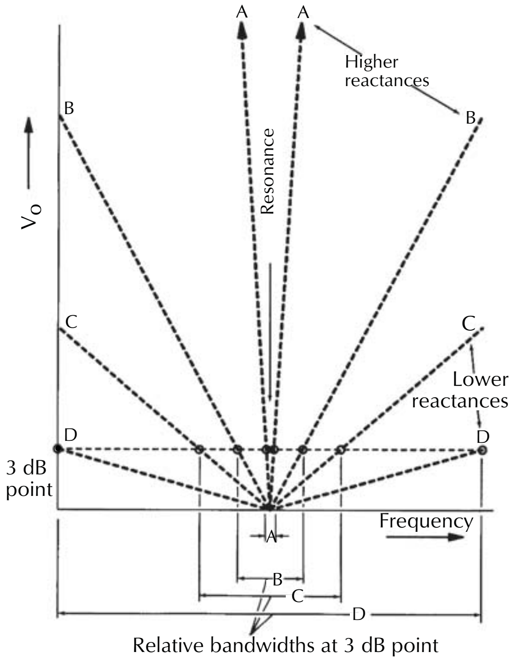


A. Series inductor and capacitor: series tuned circuit.

B. Simplified reactance plots for A.



C. Notch filter: effects shown in D.



D. Responses in C for different reactances: effect on filter bandwidth.

Figure 25-59. Series resonant circuits.

25.11.4 25.11.4 Reactance and Phase Shifts

In, for example, the context of a simple resistor/reactor filter (Fig. 25-58A), the reactance not only causes an amplitude shift with frequency but also a related phase shift. A fundamental difference between the two types of reactance (*C* and *L*) is the direction of the output voltage (V_o) phase shift with respect to the source (V_{in}). More specifically, the capacitor in Fig. 25-58A causes the output voltage phase to lag farther behind the input as the roll-off progressively bites to a limit of −90° at the maximum roll-off of the curve, while the inductor of Fig. 25-58C imposes an increasing voltage phase-lead as the low-frequency roll-off descends with a limit of +90° at maximum attenuation.

The two reactances, in their pure forms, effect phase shifts of +90° to −90° to an ultimate extent of 180° opposed; they are in exact opposition and out of phase with each other.

Referring again to Fig. 25-59, a slightly different light shines. The two reactances are working in direct opposition to each other with the inductive reactance trying to cancel the capacitative reactance and vice versa. Arithmetically, it is surprisingly simple with the two opposing reactance values directly subtracting from each other; the combination network behaves as a single reactance of the same reactive character as the one predominant in the network.

For example, if for a given frequency, the inductive reactance is a +1.2 kΩ (the + indicating the phase shift character of inductance) and the capacitive reactance is −1.5 kΩ, then the effective reactance of the entire network is that of a capacitor of −300 Ω reactance.

25.11.5 Resonance

Resonance is the strange state where the reactances of both the *L* and *C* are equal. For any inductor-capacitor pair at resonance, the two reactances will be equal. If you subtract two equal numbers, the answer is zero. So, for the series tuned-circuit arrangement of Fig. 25-59A at resonance, there is no impedance. The two reactances have canceled themselves out. It is a short circuit at that one frequency of resonance, disallowing component losses and is, in effect, a frequency-selective short circuit. Either side of that frequency, of course, one or the other of the reactances becomes predominant again.

25.11.6 Resonant Q

Like the single-order networks, there is an infinite number of combinations of *C* and *L* at any given frequency that will achieve resonance (i.e., the two reactances are equal). Similarly, it is the scale of impedance that alters with such value changes; the magnitude and rate of change of reactance on either side of resonance (off tune) hinges on the chosen combination.

At resonance, although the two reactances negate each other, they both still individually have their original values. Off resonance, their actual reactances matter. If each of the reactances is 400 Ω at resonance, then 10% off tune either way they are going to become 440 Ω and 360 Ω, respectively. A 10% change in this instance equates to about a 40 Ω change either way, up or down. Now imagine that a smaller capacitor and a larger inductor were used to obtain the same resonant frequency. Their reactances will be correspondingly larger. If they're five times larger with reactances of 2 kΩ each, then at 10% off tune their reactances will become 2.2 kΩ and 1.8 kΩ or 200 Ω change each. The higher the network impedance, the more dramatic the reactance shift off tune.

On its own, the series-tuned circuit with whatever impedances are involved doesn't amount to much; however, in relation to the outside world, it becomes rather exciting. In Fig. 25-59C the series-tuned circuit is fed via a series resistor with the output being sensed across the tuned circuit. Fig. 25-58D shows input-output curves for three different tuned-circuit impedances based on low, medium, and high reactances with the series resistor kept the same in all cases. The detune slopes are steeper with higher reactance networks than with lower ones. In other words, higher reactance networks have a sharper notch filter effect, less bandwidth, and are said to have a higher *Q* (quality factor) than lower reactance networks. In all cases, the output sensed voltage would be the same as measured across a single reactance of the appropriate and predominant sort; there is no magic about a series-tuned circuit other than the curious subtractive behavior of the two reactances.

25.11.7 Bandwidth and Q

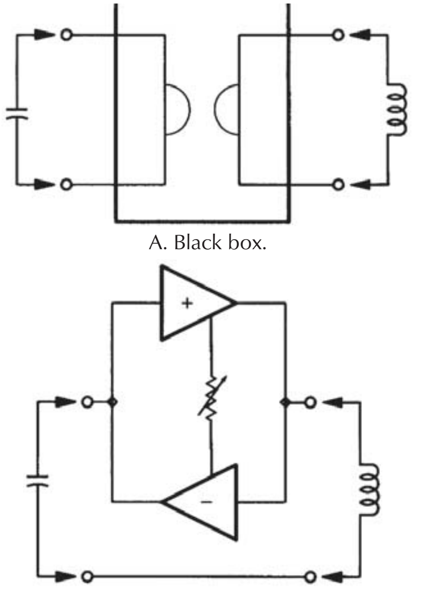
There are direct relationships between the network reactances, the series resistance, the bandwidth, and the *Q*. *Q* is numerically equal to the ratio of elemental reactance to series resistance in a series-tuned circuit ($Q = X/R$); on a more practical level, the *Q* can also be determined as the ratio of filter center frequency to bandwidth ($Q = f/\mathrm{BW}$). Bandwidth is measured between the 3 dB down points on either side of resonance (and usually where the phase has been shifted ±45°). If a tuned circuit has a center frequency of 1 kHz and 3 dB down points at 900 Hz and 1.1 kHz (pedantically 905 Hz and 1.105 kHz), the bandwidth is 200 Hz and the network *Q* is 5 (frequency/bandwidth). The greater the *Q*, the smaller the bandwidth.

The filter resonant frequency may be altered by changing either the inductance or capacitance. *Q* is subject to variation of the resistor or simultaneously juggling the reactances in the inductance-capacitance network, while maintaining the same center frequency.

25.11.8 Creating Inductance

It is most efficient (electrically and financially) in the majority of console-type circuitry for inductance to be simulated or generated artificially by circuits that are the practical implementation of a mathematical conjuring trick. These are known generically as gyrators.

A true gyrator is a four-terminal device that transmutes any reactance or impedance presented at one port into a mirror image form at the other port, Fig. 25-60A.


A. Black box.

B. Using opposed transconductance amplifiers.

Figure 25-60. Gyrators.

A capacitor on the input (with its falling reactance versus frequency) creates inductance (with a rising reactance versus frequency) at the output port. The scale of inductive reactance generated may be easily and continuously varied by altering the internal gain-balance structure of the gyrator in Fig. 25-60B by changing the transconductance of the back-to-back amplifiers, creating a continuously variable inductor.

Real inductors have a justifiably bad name for audio design, sharing transformers' less pretty attributes. They are big and heavy and they saturate easily. Their core hysteresis causes distortion, and they are prone to pickup of nearby electromagnetic fields (principally power line ac hum and RF unless well screened, which makes them even bigger and heavier). The windings and terminations are prone to break. And they are expensive.

It is quite easy to see why it is popular to avoid using real inductors. Naturally, the simulated inductive reactance is only as good as the quality of the capacitative reactance it is modeled upon and the loading effect of the gyrator circuit itself. Degradation of the inductance takes the general form of effective series lossy resistance, the Q of the inductors suffering ($Q = X/R$). Leakage resistance across or through the image capacitor is partially to blame here. Fortunately, for the purposes of normal equalizers, very large Qs are neither necessary nor desirable, so selecting capacitor types to this particular end is hardly necessary.

An obvious extension of the continuously variable inductor is the continuously variable bandpass filter formed by adding a capacitor either in series or parallel with the gyrated inductor, forming series- and parallel-tuned circuits to make notch and peak filters, respectively. Although ideal for fixed-frequency filters with the Q of the network or sharpness defined by a resistor in series with the gyrator resonator, the idea falls down when the resonance frequency is moved.

If the frequency is moved higher by altering either the L or C, the reactances of the element at resonance become lower; consequently, the ratio of the reactances to the fixed-series resistor (this is the ratio that determines the Q) becomes smaller, and the bandwidth of the filter becomes broader in response. In order to maintain the same Q over the projected frequency variation, the series resistor has to be ganged with the frequency control, which is not easy. Should it be necessary to make the Q a variable parameter also, as in a parametric-type EQ section, it would mean devising quite a complex set of interactive variable controls. For this reason parametric-type EQ sections are ordinarily constructed around second-order, active-filter net-works, not individual tuned circuits whether real or gyrated.

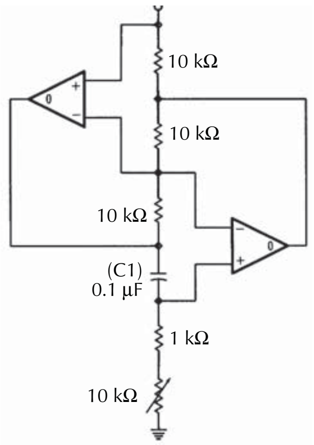
25.11.9 Gyrator Types

Let us not write off gyration for functionally variable filters immediately. As we'll see, they form in one way or another the second reactance in many active filters.

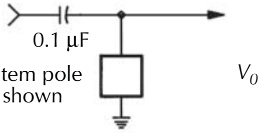
True gyrators of the back-to-back transconductance amplifier variety are difficult to make, set up, and use. Fortunately, there are simpler ways of simulating variable reactances—if not pure reactances at least a predictable effect of a reactive/resistive network.

25.11.10 The Totem Pole

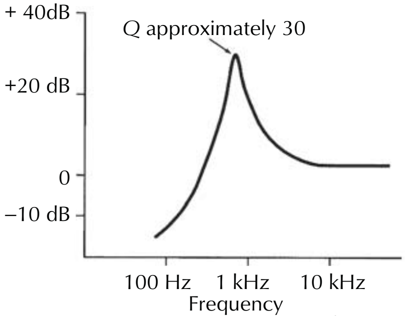
Fig. 25-61 performs the magic transformation of the single capacitor C1 into a simulated inductance between the terminals. Although emulating quite a pure inductance when set up properly, it is precisely that setting up that is not altogether straightforward. In fact, it is high on a list of circuits most likely to do undesired things.


10 kΩ
10 kΩ
10 kΩ
(C1)
0.1 µF
1 kΩ
10 kΩ
0
+
−

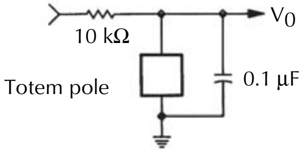
A. Totem pole gyrator with typical values.


0.1 µF
Totem pole
as shown
V_0

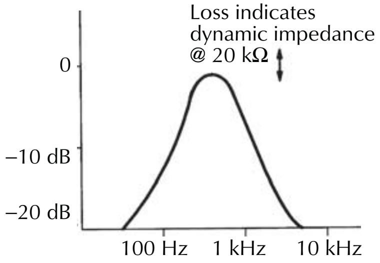
B. High-pass filter.



C. Input output response of B.


10 kΩ
V_0
Totem pole
0.1 µF

D. Bandpass filter.

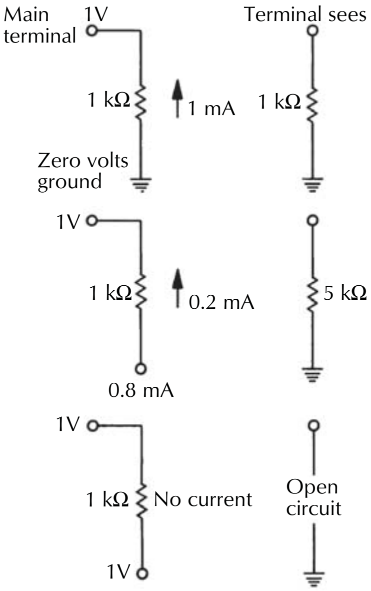


E. Input/output response of D.

Figure 25-61. Totem-pole gyrators.

25.11.11 The Bootstrap

The simplest fake inductor is shown in Fig. 25-63A, with typical values. It relies on a technique called *bootstrapping*. The principles are shown in Fig. 25-62. A 1 kΩ resistor with 1 V across it will pass 1 mA. Without changing the source potential of 1 V, the bottom end of the resistor is tied to 0.8 V. There is 0.2 V across the resistor, and so a current of 0.2 mA flows through the resistor. The source (still at 1 V) sees 0.2 mA flowing away from it, the amount of current it would expect to see going to a 5 kΩ resistor value (1 V/0.2 mA = 5 kΩ). It thinks it's looking at a 5 kΩ resistor! Continuing this, stuffing a potential of 1 V (not the same source) at the bottom end of the resistor means there is no voltage across the resistor, so there is no current flow. Our original source thinks it's seeing an open circuit (infinite resistance) despite the fact that there is still a definite, real, physical 1 kΩ resistor hanging on it.

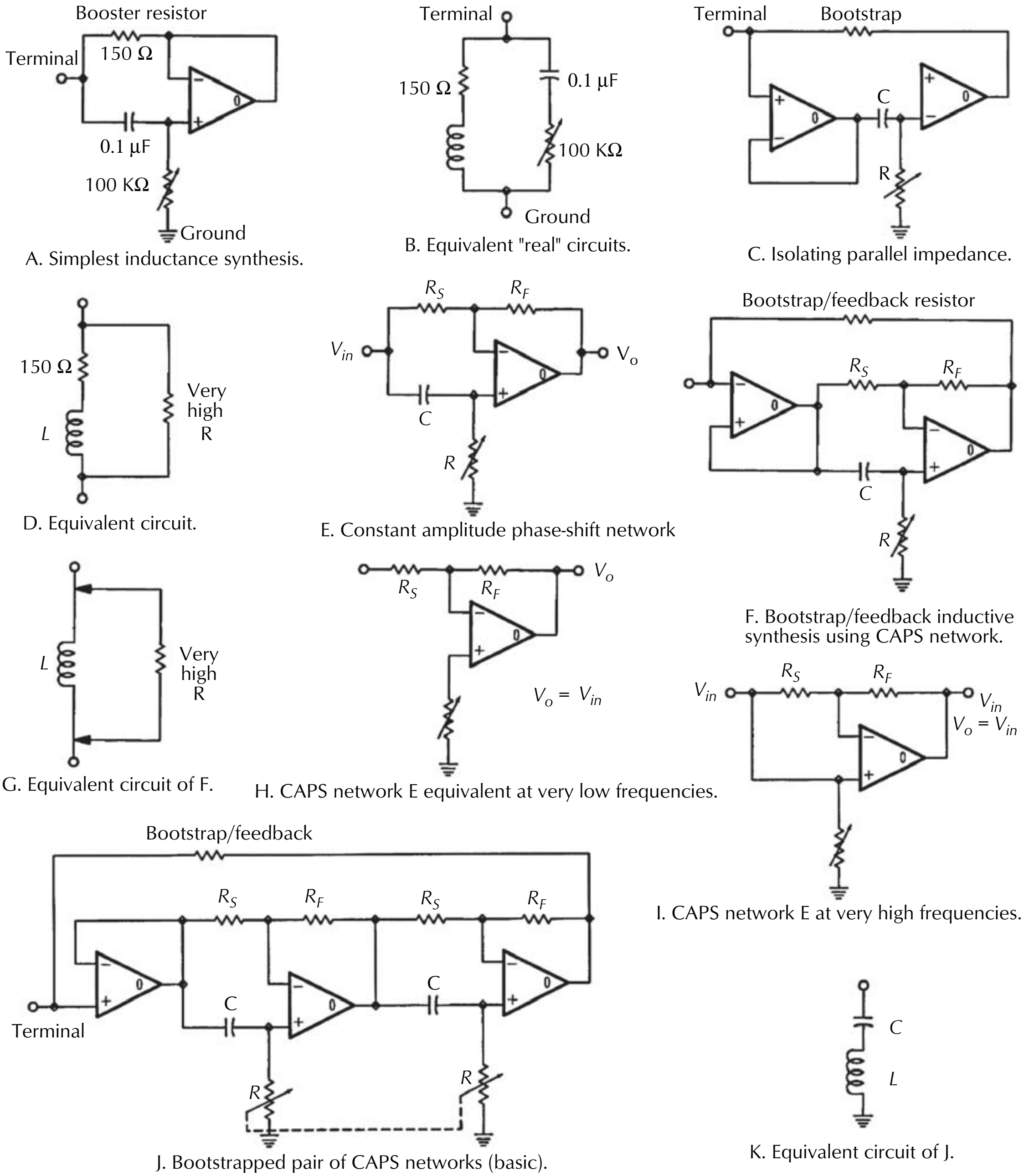

Main
terminal
1V
Terminal sees
1 kΩ
1 mA
1 kΩ
Zero volts
ground
1V
1 kΩ
0.2 mA
5 kΩ
0.8 mA
1V
1 kΩ
No current
Open
circuit
1V

Note: 1 kΩ resistor is attached
to terminal in all cases

Figure 25-62. Bootstrapping analysis.

This phenomenon holds true with any source voltage, ac or dc, provided the instantaneous bootstrap voltage is the same as the source. Any phase or potential difference creates an instantaneous potential difference across the resistor; current flows and an apparent resistance materializes.

This fake inductor works on frequency-dependent bootstrapping, the terminal being almost totally boot-


Booster resistor
Terminal
150 Ω
0.1 µF
100 KΩ
Ground
Terminal
150 Ω
0.1 µF
100 KΩ
Ground
Terminal
Bootstrap
C
R
150 Ω
L
Very high R
R_S
R_F
V_{in}
V_o
C
R
Bootstrap/feedback resistor
R_S
R_F
C
R
L
Very high R
R_S
R_F
V_o
$V_o = V_{in}$
V_{in}
R_S
R_F
V_{in}
$V_o = V_{in}$
Bootstrap/feedback
R_S
R_F
R_S
R_F
C
C
R
R
Terminal
C
L

A. Simplest inductance synthesis.

B. Equivalent "real" circuits.

C. Isolating parallel impedance.

D. Equivalent circuit.

E. Constant amplitude phase-shift network

F. Bootstrap/feedback inductive synthesis using CAPS network.

G. Equivalent circuit of F.

H. CAPS network E equivalent at very low frequencies.

I. CAPS network E at very high frequencies.

J. Bootstrapped pair of CAPS networks (basic).

K. Equivalent circuit of J.

Figure 25-63. Inductive reactance synthesis.

strapped to high impedance via the 150 Ω resistor at high frequencies and the bootstrap voltage reducing (together with its phase being shifted) with falling frequency. At very low frequencies the capacitor behaves as a virtual open circuit. No bootstrap exists, so the terminal is tied to ground via the 150 Ω resistor and the effectively zero output impedance of the voltage follower. The circuit emulates an inductor reasonably well; it has a low-impedance value at low frequencies, increasing with frequency to a relatively high impedance.

A minor failing with this simple circuit is that at high frequencies a parallel impedance (consisting of the variable resistor and capacitor chain) hangs directly from the terminal to ground. Buffering the chain from the terminal by a follower eliminates this, Fig. 25-63C.

Fig. 25-63A creates an analog of an inductor with the losses shown in Fig. 25-63B. The series resistor is the 150 Ω bootstrap resistor; after all, a proper inductive reactance tends to zero at low frequencies, not 150 Ω. The resistor is in series with the faked inductance tending to make it seem somewhat lossy or have a lower *Q* than a perfect inductor. If a fake inductor can be said to have winding resistance, this is it! The *R*/*C* network across the lot represents, again, the high-pass filter impedance, which on the addition of the follower disappears to be replaced in Fig. 25-63D by the much greater input impedance of the follower, which is high enough to be discounted.

As a short footnote to this gyrator epic, consider what happens to either Fig. 25-63C or F if the high-pass resistance-capacitance filter is replaced by a low-pass filter by swapping *R* with *C*. It may seem a bit strange to use circuitry to imitate a capacitor, but imitating a continuously variable capacitor does make sense. Real variable capacitors of the large values needed in EQs (yet easily created by gyrators) simply don't exist otherwise.

25.11.12 Constant-Amplitude Phase-Shift Network

A constant-amplitude phase-shift (CAPS) circuit of previously little real worth (other than for very short time delays) is shown in Fig. 25-63E. Bearing more than a little resemblance to a differential amplifier, this circuit can rotate the output phase through 180° with respect to the input, around the frequency primarily determined by the high-pass RC filter. Additionally the input and output amplitude relationship remains constant throughout.

How? This is dealt with in Figs. 25-63G and H where the simplistic assumptions that a capacitor is open circuit at low frequencies and a short at high frequencies show that at low frequencies the circuit operates as a straightforward unity-gain inverting amplifier (−180° phase shift), while at high frequencies it operates as a unity-gain noninverting amplifier (0° shift). The mechanism for the latter mode is interesting. The op-amp is actually operating at gain-of-two noninverting; this is compensated for by the input leg also passing through the still operating unity inverting path, which naturally subtracts to leave unity gain, noninverting.

25.11.13 Simulated Resonance

Detailed up to here are all the variables needed to create single- and second-order filters. Higher-order networks can be made with combinations of the two. Tracking variable capacitors and inductors allows the design of consistent *Q* bandpass filters irrespective of frequency. This eventually leads to a dawning of understanding in how the much-touted integrator-loop filters such as the state variable actually operate. The clue lies with the 180° phase-shift circuit of Fig. 25-63E. Connecting two such filters (with the variable resistor elements ganged) in series produces a remarkably performing circuit. At any frequency within the design swing, it is possible for the circuit output voltage to be exactly out of phase with the source −180° phase shift). By summing input and output, direct cancellation at that frequency and at no other is achieved. In short, a variable-frequency notch filter with a consistent resonant characteristic results. Alternatively, bootstrapping the input from the output actually changes that input port into something that behaves exactly like a series-tuned circuit to ground, Fig. 25-63J. The circuit is continuously variable in frequency with a consistent *Q* by virtue of the simultaneously tracking simulated inductor and capacitor maintaining exactly the same elemental reactances at whatever the selected operating resonant frequency. This creates the same source resistance, same reactance, same *Q*.

25.11.14 Consistent Q *and Constant Bandwidth*

The same *Q* definitely does not imply the same filter bandwidth. As the resonant frequency changes, the bandwidth changes proportionally. Bandwidth is, after all, the ratio of frequency to *Q*. Some active filters, such as the multifeedback variety, exhibit a constant bandwidth when the resonant frequency is changed: a 10:1 variation of center frequency, a 10:1 variation of *Q*. This, of course, is rarely useful for real EQ; it is noteworthy though in that the change in *Q* with frequency happens in the opposite sense to that expected from a normal variable tuned circuit. The *Q* sharpens with increasing frequency. It is a perfect example of a constant-bandwidth filter.

25.11.15 Q, EQ, and Music

The near insistence on resonant-type filters being constant in *Q* when varied in frequency is not through an industrywide collective lack of imagination or desire to keep things tidy. It stems from psychoacoustics, from the way humans react to audible stimuli, and also from the way nature deals with things acoustic.

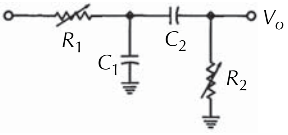
If something is acoustically resonant, it will need a similar electrical resonance response shape to compensate for, extract, or imitate it in the console. Acoustics are defined by exact analogs of the first- and second-order filters and the time-domain effects that we've been delving into here in EQ.

Differing wall coverings have absorption coefficients paralleled very closely with shelving-type EQ curves. Apertures and partly enclosed spaces big and small act like second-order resonances identical to electrically resonant circuits. The physical size of a room determines the lowest frequency it can support just as a high-pass filter would. Initial and other major room reflections effect precisely the same changes on audio as deliberately introduced electronic delays; the frequency-dependent propagation characteristics of air are emulatable with slope filters.

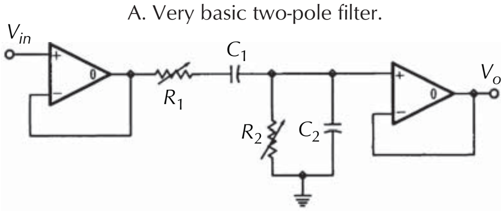
25.11.16 Bandpass Filter Development

Methods of filtering come thick and fast once the basics are established. The development of a popular bandpass filter arrangement is shown in Fig. 25-64. It starts as two variable passive single-order filters of a common crossover frequency point, ganged so that they track. Reconfigured slightly, Fig. 25-64B, to minimize interaction, they are shown with their drive and sense amplifiers. Wrapping the two networks around an inverting amp isolates them completely from each other, improving the filter shape. The bandpass *Q* is rather low, well under one, leaving it rather limited in scope for practical applications. Positive feedback from the amplifier output back to the noninverting input sharpens the *Q*.

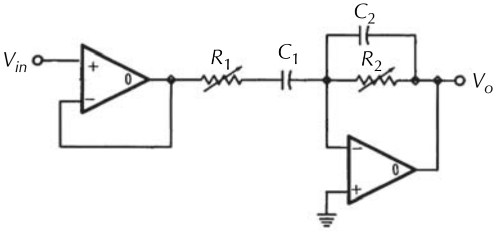
Yes, it does look rather like a Wein Bridge oscillator. Attempting to get the *Q* too high proves the point unquestionably!



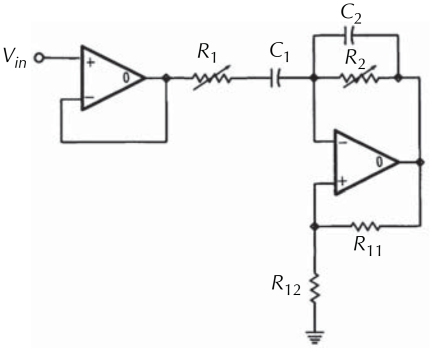
A. Very basic two-pole filter.



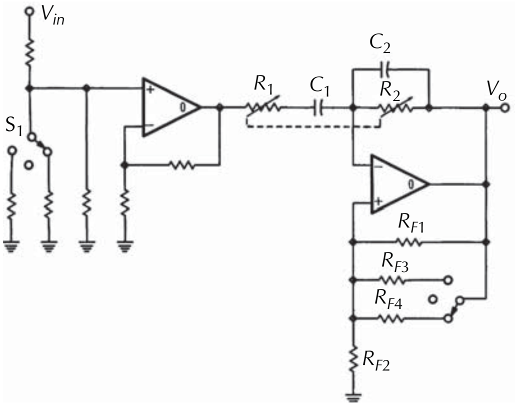
B. Reconfigured with source and sense amplifiers.



C. Elements isolated around inverting amplifiers.



D. Positive feedback introduced to increase Q.



E. Switchable Q with ganged switched compensating attenuator.

Figure 25-64. Bandpass filter development.

25.11.17 Listening to Q

This raises the problems of excessive *Q*s. Fortunately, extremely high *Q*s (greater than ten) are unnecessary or unusable for EQ purposes. The higher the *Q* becomes, the less actual spectral content of the signal it modifies, so despite the fact that its peak gain or attenuation is the same as a lower *Q* filter, it seems to do subjectively less. Judicious care is required in setting up the filter to enhance or trim exactly what is required. Accidental overkill is easy.

There comes a breakpoint with increasing Q where you are not so much listening to the effect of the filter as to the filter itself. Resonant-tuned circuits are essentially ac electrical storage mechanisms, where energy inside the circuit shuffles backward and forward between the two reactive elements until the circuit losses waste it away. The greater the Q (and by definition the lower the included losses), the more pronounced this signal storage is.

Think of a high-Q circuit as a bell, which is just an acoustic version of the same thing. If the bell gets booted either physically or by being excited by audible frequencies at its tuned pitch, it will ring until its natural decay. It's the same with a filter. A transient will set it ringing with a decay time related to the filter Q. Music containing energy at the filter frequency will set it off just as well; a listener will hear the filter ringing long after the original transient or stimulus has stopped. Despite being good for a laugh, extremely high Qs and the resultant pings trailing off into the sunset are of no value whatsoever in a practical EQ. A transient hitting such a filter fires off a virtually identical series of decaying sine waves at the frequency of the filter.

Square waves sent through audio paths are good for kicking resonant ringing off at almost any frequency. It's a convenient means of unearthing inadvertent response bumps, phase problems, and instabilities. The breakpoint—where filter ringing is as audible as signal—is quite low, a Q of between five and ten depending on the nature of the program material.

25.11.18 Push or Retard?

It is not too difficult now to appreciate that resonant circuits and oscillators are very close cousins—often indistinguishable, except for maybe an odd component value here and there. There are two fundamental approaches to achieving a resonant bandpass characteristic using active-filter techniques.

The first is to start off with a tame, poorly performing, passive network and then introduce positive feedback to make it predictably (we hope) unstable. The feedback exaggerates the filter character and increases the Q to the desired extent. A perfect example of this is the Wein Bridge development of Fig. 25-64. The major disadvantage of such methods is that the Q is disproportionately critical with respect to the feedback adjustments, especially if tight Qs are attempted.

The second approach is to start off with an oscillator and then retard it until it's tame enough. This is the basis of the state variable, the biquad, and similar related integrator-loop-type active filters.

25.11.19 The Two-Integrator Loop

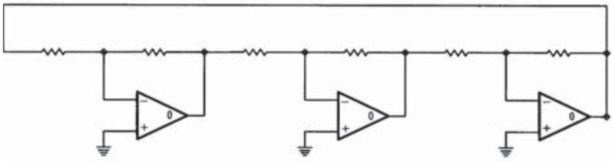
This, for better or worse, and a variety of reasons, is by far and away the most popular filter topology used in parametric equalizers. Three inverting amplifiers connected in a loop, as shown in Fig. 25-65, seem a perfectly worthless circuit and, as such, it is. It's there to demonstrate (assuming perfect op-amps) that it is a perfectly stable arrangement. Each stage inverts (180° phase shift), so the first amplifier section receives a perfectly out-of-phase (invert, revert, invert) feedback, canceling any tendency within the loop to drift or wobble. Removing 180° phase shift would result in perfect in-phase positive feedback; the result is an oscillator of unknown frequency determined predominantly by the combined propagation times of the amplifiers.

Arranging for the 180° to be lost only at one specific frequency results in the circuit being rendered unstable at just that one frequency. In other words, it oscillates controllably. Creating the 180° phase loss is left to two of the inverting amps being made into integrators, Fig. 25-65B, so called because they behave as an electrical analog of the mathematical function of integration.

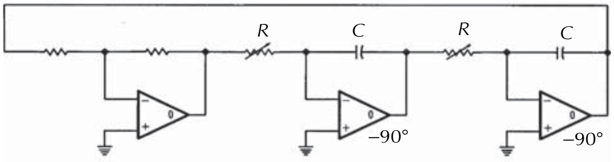
The integrator you may recognize from a single-order filter variation in Fig. 25-59. It's not so much the amplitude response that's useful here as the phase response, which at a given frequency (dictated by the R and C values) reaches –90° with respect to its input. Two successive ganged-value integrators create a 180° shift.

Retarding the loop to stop it from oscillating can be achieved in a variety of ways:

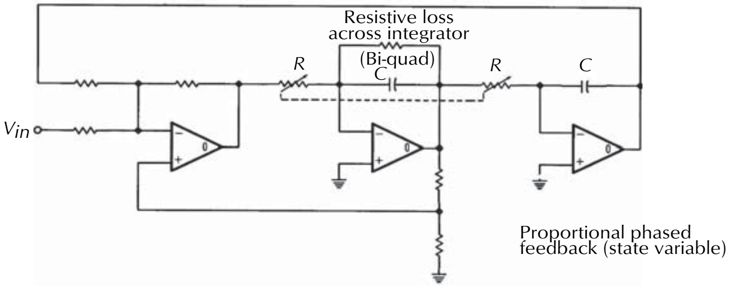
1. Trimming the gain of the remaining inverter. This is unduly critical like the Wein Bridge for Q determination.
2. Doping one of the integrator capacitors with a resistor, Fig. 25-65C. This in essence is the biquad filter (after *biquadratic*, its mathematical determination). The Q is largely dependent on the ratio of the capacitive reactance to the parallel resistance; consequently, it varies proportionally with frequency. For fixed-frequency applications the biquad is easy, docile, and predictable.
3. Phased negative feedback. This is not true negative feedback but taken from the output of the first integrator (90° phase shift). It provides an easily managed Q variation, is constant, and is independent of filter frequency, Fig. 25-65C. Forming the basis of the state-variable filter, this has turned out to be "the active filter most likely to succeed," if the majority of current commercial analog console designs are to be believed.



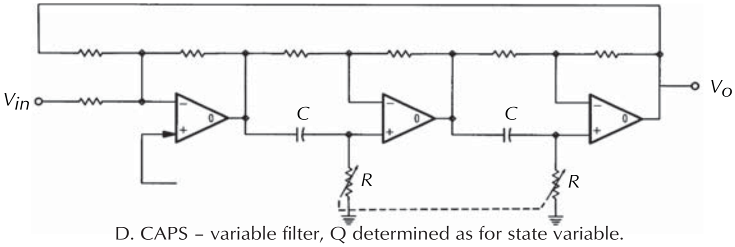
A. Three-inverter loop (stable).



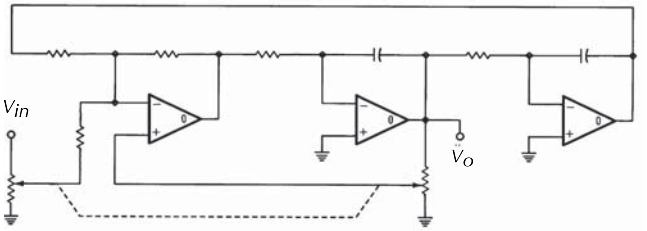
B. Introducing 180° phase shift via two integrators.


Resistive loss
across integrator
(Bi-quad)
R
C
R
C
V_{in}
Proportional phased
feedback (state variable)

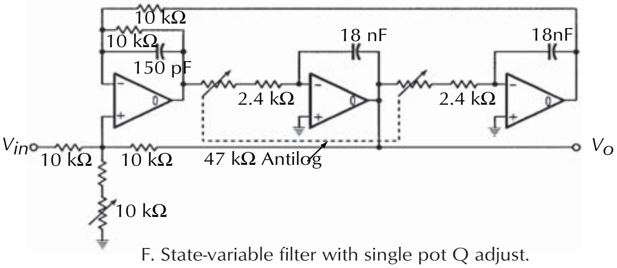
C. Introducing loss to temper the Q to usability.



D. CAPS – variable filter, Q determined as for state variable.



E. Ganged potentiometer simultaneously altering Q and input attenuator.


10 kΩ
10 kΩ
150 pF
18 nF
18nF
2.4 kΩ
2.4 kΩ
V_{in}
10 kΩ
10 kΩ
47 kΩ Antilog
V_O
10 kΩ

F. State-variable filter with single pot Q adjust.

Figure 25-65. Loop filters.

Loop filters, such as described in Fig. 25-65, have a number of inherent problems that are usually glossed over for the sake of the operational simplicity and elegance of the design.

25.11.20 Stability and Noise Characteristics

Each amplifier within the loop has a finite time delay, which together add up to significant phase shifts within the open loop bandwidths of the amplifiers. Some simply add to the delay imparted by the integrators, but the total time discontinuity around the summing amp can promote instability in the multimegahertz region. Compensation for this around the summing amplifier can introduce further phase shifts, upsetting the filter performance at high frequencies.

Two major problems are due to the nature of the integrator arrangement itself. They come to light at the extremes of the feedback capacitive reactance (i.e., at very low and very high frequencies where, respectively, the reactances are virtually open circuit and short circuit).

Open circuit at low frequencies means the op-amp is infinitely amplifying external resistor noise and internally generated thermal and (mostly) low-frequency 1/F noise, plus any low-frequency noise presented to the input along with the signal. There is a lot of generated and circulating low-frequency noise.

At high frequencies, the reactance approaches a short circuit, connecting the output back around to the inverting input. This arrangement, zero closed loop gain, is about as critical in terms of device instability as it can get. It is directly analogous to a grounded-input follower (see Section 25.7 for inherent problems), since there is no possible way of further externally defining the closed loop characteristics beyond those of the integrating capacitor itself. For typical audio frequency EQ the integration capacitor value can be quite sizable, up to 1 µF. Two further aggravations:

1. Current limiting. Is the current output capability of the op-amps sufficient to charge such a size capacitor instantaneously? If not, this will result in low maxima of signal frequency and signal level before op-amp slew-rate limitation sets in. The amplifier just might not be able to deliver enough current quickly enough.
2. Finite device output impedance. There will almost certainly be another foible related to the open loop output impedance of the op-amp; this corresponds to a resistor in series with the device output that forms a time constant and a filter with the integrator capacitor, in addition to the intended one. Another time constant means more time delay in the loop, causing a seriously degraded (maybe already critical) stability phase margin. At best it adds a zero to the integrator, reducing the integrator's effectiveness at high frequencies.

Integrators ask a lot of device outputs; not only do they have to cope with a vicious reactive load (with which many op-amps are ill equipped to cope) but they also have to drive other circuitry, such as the next stage. A mad drive to bring circuit impedances down for noise considerations can soon outstrip even the best op-amp's capabilities.

As tame as it may superficially seem, the state variable is not an unconditionally or reliably stable arrangement, with out-of-band dynamic problems potentially degrading its sonic performance. It is an amazement that these filters work as well as they do in many commercial designs.

With the exception of inevitable loop effects (usually time related), most of the undesirable things about the state variable can be eliminated or mitigated by replacing the integrators with constant amplitude, phase-shift elements, Fig. 25-65D. This results in what could best be known as a CAPS-variable filter. Here, all the constituent elements are basically stable, and there are provisions for independent device compensation. There is no undefined gain for any of the spectrum. This seems to be a far healthier format to start making filters around.

There is another way of looking at the state variable/CAPS-variable filters that will suddenly resolve the previous discussions on gyrators, *L* and *C* filters, series-tuned circuits, and so on with the seemingly at-odds approach of active filters.

Resonance depends on the reaction of the two reactances of opposite sense, 180° apart in phase effect. Rather than achieve this in a differential manner, one element +90° with the other –90° at a given frequency, these active filters achieve the total difference by summing same-sense phase shifts (–90° + –90°)–i.e., still 180° apart. Two reactive networks are still involved; it is still a second-order effect. At the end of the day, the principal difference is that such loop-type active filters have their median resonance phase displaced by 90° from their input as a result of both reactive effects being in the same sense, as opposed to the nil phase shift at resonance of a real *LC* network.

25.11.21 Q and Filter Gain

Pretty much every resonant-type active filter has the characteristic of its gain at resonance being at least related and often directly proportional numerically to the *Q* of the filter. This means a filter with a *Q* of 10 usually has a voltage gain of 10 (20 dB) at resonance. Naturally, this does not make the building of practical equalizers any easier. Nothing much does. Even specifying a maximum *Q* of 5 (14 dB gain) only helps by losing 6 dB of boost with respect to a *Q* of 10.

That represents a very sizable chunk of system head room stolen at the filter frequency, which also makes the sum-and-difference matrixing necessary to provide the usual boost-and-cut facilities difficult to configure. The obvious solution is to attenuate the signal going into the filter by the same amount as the gain and *Q* expected of the filter. Arranging a continuously variable *Q* control that also attenuates the signal source appropriately is not a conspicuously simple task, at least with most filters. Perhaps the most straightforward example is shown in Fig. 25-65C, a state-variable-type filter with an attenuator in the retard network altering the *Q* ganged with an attenuator ahead of the input/summing amplifier. Within reasonable limits this holds the resonant peak output constant over a considerably useful *Q* range. A much neater and more commonly applied solution is shown in Fig. 25-65F: a single potentiometer at the noninverting input of the summing amp that would serve both purposes—filter *Q* and input level—complementarily and simply.

Most other filters are not so obliging in terms of continuously variable *Q*. Switching between a few values of *Q* while substituting appropriate input attenuation is quite often a practical and operationally acceptable solution, applicable to nearly any filtering technique. Fig. 25-64E illustrates a further development of the Wein Bridge arrangement using this method to provide three alternative *Q*s. The attenuator values are necessarily high in impedance to prevent excessive loading of the source, a factor that in some practical EQ circumstances can be important.

25.11.22 High-Pass Filters

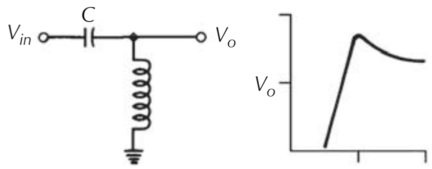
Two basic single-order high-pass filters are shown in Fig. 25-66. The keys, for the purposes of high-pass filtering, are the reduction of inductive reactance to ground with reducing frequency in Fig. 25-66F and the increasing of capacitative reactance with reducing frequency in Fig. 25-66G.

How about combining the two and omitting the resistors as in Fig. 25-66A? As expected, the combining of the two opposing reactances causes an ultimate roll-off twice as fast as for the single orders; however, they have also resulted in a resonance peak at the point of equal reactance. Resonance *Q* is the ratio of elemental reactance to resistance; deliberately introducing loss in the circuit in the form of a termination resistor tames the resonance to leave a nice, flat, in-band response, Fig. 25-66B.

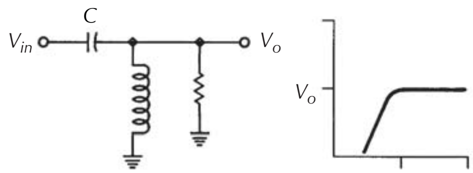
Substituting a basic gyrator or simulated inductance for the real one, Fig. 25-66C naturally works just as well and even better than expected. The filter output can be taken straight from the gyrator amplifier output, eliminating the need to use another amplifier as an output buffer. Further, we can automatically introduce the required amount of loss into the inductor by increasing the value of the bootstrap resistor and get the resonance damping right. (Refer to the discussion of gyrators in Section 25.11.9.)

Further yet, we can easily change the turnover frequency of the filter by varying what was the tuning resistor. In doing this, of course, the elemental reactance-to-loss ratio will change, causing damping factor (and so the *Q*) to change with it. The frequency change and required damping change are directly related and in the same sense and may be simultaneously altered with a ganged control—even, if we do our sums right, with the two ganged tracks having the same value!

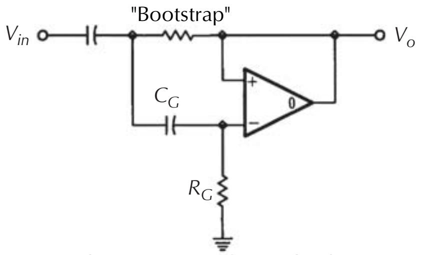
A slight redraw of Fig. 25-66C gives Fig. 25-66D, a more conventional portrayal of the classic Sallen-Key high-pass filter arrangement. As the Sallen-Key filter evolves, it turns out that an equal value filter (where the two capacitors are equal and the two resistors are equal) results in a less than adequate response shape. An expedient method of tailoring and smartening up response to become Butterworth-like (working on the assumption that a couple more resistors are cheaper than a special two-value ganged potentiometer) is to alter the damping by introducing gain into the gyrator buffer amplifier (providing also a healthier mode of operation for the amplifier—followers are bad news), see Fig. 25-66E. A side effect of this technique of damping adjustment (which, incidentally, is independent of filter frequency) is that an input-output in-band gain is introduced. The 4 dB gain introduced necessary to render the filter frequency response maximally flat could be included in overall system gain, or alternatively a compensating attenuator could be instituted ahead of it. This could as well be arranged to be a fixed-frequency, band-end, single-order, high-pass filter to accelerate the roll-off slope out of band; a further alternative is to make the



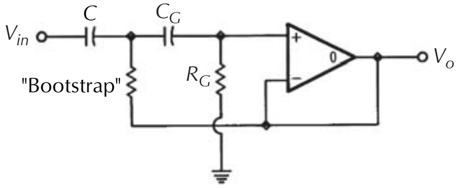
A. A simple inductance-capacitance high-pass filter.



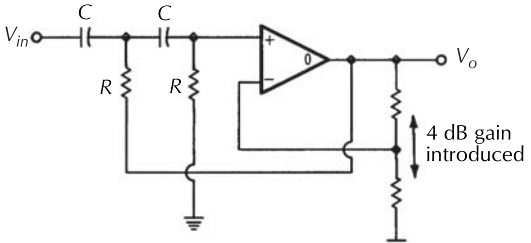
B. High-pass filter, with critical damping.



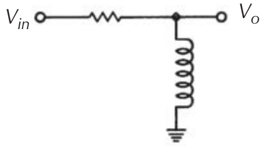
C. Substituting a synthesized inductance.



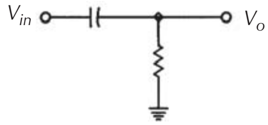
D. Circuit C redrawn revealing classic Sallen-key high-pass filter.



E. Introducing gain, enabling conveniently equal values of resistance and capacitance to be used.



F. Simple high-pass resistance-inductance filter.



G. Simple high-pass resistance-capacitance filter.

Figure 25-66. High-pass filter development.

filter input out of two capacitors such that the input signal is attenuated by the needed amount yet the combined capacitance value is the correct value for the filter—this can be a bit of a nightmare to drive adequately, though. For many applications the free 4 dB or so isn't a problem—it can simply be assimilated as part of the system-level architecture.

The 4 dB thing can be a nuisance. Where it is, or in particular where an inverting filter stage is either convenient or necessary, the multifeedback configuration works well; indeed, lacking the problems of a near-follower as in the case of the Sallen-Key it uses the op-amp well. At high values of Q or extremes of frequency, some component values can get far from the ordinary midimpedance values seen elsewhere in the EQs and filters described here, and one should be aware of possible noise or op-amp current-drive limitation issues as a consequence. Unlike the Sallen-Key described, it is not readily possible to continuously vary the turnover frequency, and it uses three capacitors as frequency-determining components rather than two. Otherwise, for fixed frequency filters this is a very friendly topology.

25.11.23 Second or Third or More Order?

Without delving too deeply into psychoacoustics, the ear notices easily third or more order filters being introduced for much the same reasons as a high-Q bandpass filter is obvious. There are severe modifications to the transient response of the signal path and ringing-type time-related components are introduced into the signal spectrum.

An application where this effect is not overly objectionable is where the filters are defining bandwidth at high and low audible extremes. Within the audible band though, the ear is quite merciless toward such artifacts.

The transient response modification and time-domain effects are not the end of the story; the relationships between instrument fundamentals and their harmonics in the turnover area of the filter are likely to be interpreted as unnatural, especially should the fundamental be attenuated with respect to the harmonics.

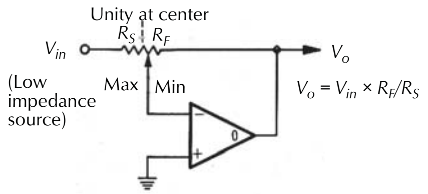
Second-order filters, assuming moderate filter Qs associated with Bessel or Butterworth characteristics, score well in both respects. There is less transient response disturbance and less tonal characteristic modification. There are few who would dispute that they sound more natural and musical than tighter filters. A small wrinkle is to leave a small controlled amount of underdamped bump in the filter frequency response. This has two consequences: one is the slightly more rapid out-of-band roll-off, but the other, a subjective effect, is that the extra program energy introduced by the hump serves to help offset the loss of energy below the turnover frequency. The perceived effect on introducing the filter is more of a slight change in sound rather than a direct drop in low-frequency response and strikes a better subjective compromise than techno-striving for the ultimately flat, perfectly measuring filter.

25.11.24 Equalization Control

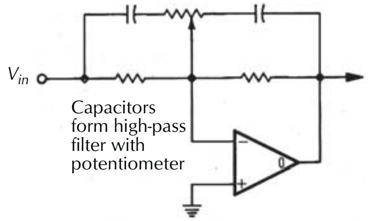
Achieving bare response shapes of whatever nature—high-pass, low-pass, bell-shaped bandpass, or notch—does not really constitute a usable EQ system. The shape, even if variable in frequency and bandwidth, is either there or not, in or out, no subtleties or shades; some means of achieving control over the strength of effect is vital to the cause. By far the most common (but certainly not the only) control requirement and one easily understood by operators is lift and cut, where the frequency areas relevant to the various filters are required to be boosted or attenuated by any variable amount within given limits. Determining these limits alone is good for an argument or two, dependent on such disparate considerations as system head room, operator maturity, and, obviously, application. An EQ created specifically for wild effects is not a stable device. An adjustment of 20dB is not unknown (and not, unfortunately, unheard); a 6 dB adjustment, in contrast, is often far more than enough particularly for spoken voice. A general median accepted by most manufacturers is to provide between ±12 and ±15 dB level adjustment on channel-type EQs.

25.11.25 The Baxandall

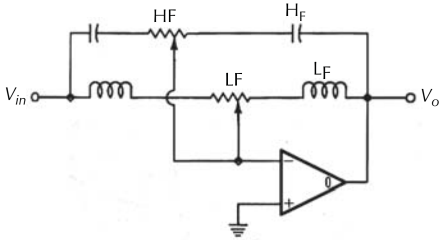
Hi-fi-type tone controls needed similar basic operational high-frequency and low-frequency boost-and- cut facilities, and a design for this dating from the 1950s by Peter Baxandall has since been an industry standard in assorted and updated forms. A development of the Baxandall idea is represented in Fig. 25-67 based


Unity at center
R_S
R_F
V_{in}
V_o
(Low impedance source)
Max
Min
$V_o = V_{in} \times R_F/R_S$
0

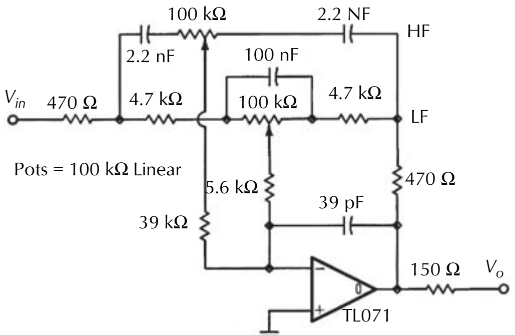
A. Infinitely variable gain-inverting amplifier.


V_{in}
Capacitors form high-pass filter with potentiometer
0

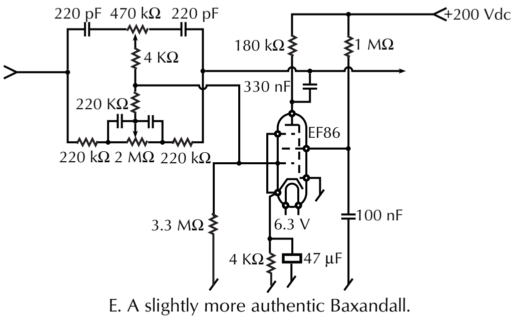
B. Continuous variable high-frequency boost and cut.


HF
H_F
LF
L_F
V_{in}
V_o
0

C. A belt-and-braces low-pass arrangement.


100 kΩ
2.2 NF
HF
2.2 nF
100 nF
V_{in}
470 Ω
4.7 kΩ
100 kΩ
4.7 kΩ
LF
Pots = 100 kΩ Linear
5.6 kΩ
470 Ω
39 pF
39 kΩ
150 Ω
V_o
0
TL071

D. An op-amp-based Baxandall-type high-frequency and low-frequency shelving equalizer.


220 pF
470 kΩ
220 pF
+200 Vdc
180 kΩ
1 MΩ
4 KΩ
330 nF
220 KΩ
EF86
220 kΩ
2 MΩ
220 kΩ
3.3 MΩ
6.3 V
100 nF
4 KΩ
47 µF

E. A slightly more authentic Baxandall.

Figure 25-67. Development of Baxandall-style equalizer.

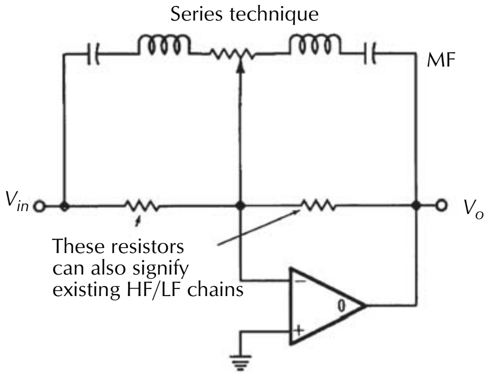
around today's more familiar op-amp technology rather than discrete transistors or tubes. Fig. 25-67A shows a virtual-earth-type inverting amplifier with the gain (being equal to the ratio of the feedback resistor RF to the series resistor RS) continuously variable from near-infinite loss (min) to near-infinite gain (max) with unity in the middle. If a fixed-gain-determining leg is introduced and the variable leg is made frequency conscious, as shown in Fig. 25-67B (in this instance by crude single-order high-pass filters—the series capacitors), the gain swing only occurs within the passband of those filters. The through gain for the rest of the spectrum is determined by the two fixed resistors. If this fixed chain is replaced by a second frequency-conscious network that does not significantly overlap the original one in bandwidth, the two chains independently modify their frequency areas, Fig. 25-67C. The fixed chain is only necessary where the gain is otherwise unpredictably defined by a frequency-conscious network.

The belt-and-braces low-pass arrangement (for low-frequency boost and cut) of Fig. 25-67C can be rationalized into the more elegant circuit of Fig. 25-67D. This circuit more closely resembles the definitive Baxandall circuit. Rather than isolating the low-frequency boost-and-cut chain with increasing inductive reactance, the control is buffered away with relatively small resistances and bypassed to high frequencies by capacitance. The control takes progressively greater effect at lower frequencies as the rising capacitative reactance reduces the effective bypass. A further refinement is a pair of stopper resistors, small in value, that define the maximum boost and cut of the entire network.

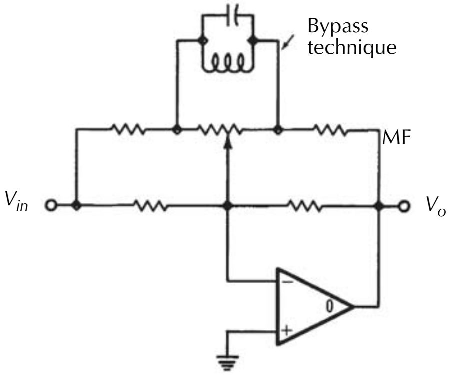
Naturally, a more complex EQ can be configured around the same arrangement. A midfrequency bell curve is easily introduced by any of the means in Fig. 25-68, giving a good hint on how to avoid using a real tuned circuit using inductors.

A variable signal either positive or negative in phase to the source V_{in} can be picked off from a pot straight across the existing high-frequency and low-frequency chains, taken to an active filter arrangement to derive the needed amplitude response shape. The signal is then returned to the loop at either the virtual-earth point (to which the high-frequency and low-frequency chains are tied) or to the noninverting reference input, Fig. 25-68D, depending on whether the absolute phase of the filter is positive or negative. Industry favorites seem to be this approach using either a Wein Bridge bandpass or a state-variable integrator-loop type.

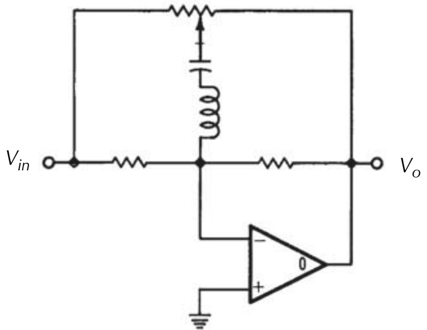
Any number of such active chains may be introduced, provided two great hangups don't intrude excessively:



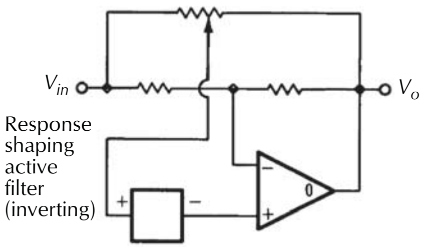
A. Series–tuned symmetrical bandpass.



B. Single parallel–tuned element.



C. Single series–tuned bypass filter.



D. Using an active filter element.

Figure 25-68. Resonant frequency selective elements in the Baxandall equalizer.

- Hangup 1 is the interaction between frequency groups. Hanging on two control chains that operate at the same frequency either adjustably or through overlap can at best be deceiving or at worst self-defeating. In the Baxandall (as with most other arrangements), if maximum gain (say, 15 dB) is attained at a given frequency by one control, a second similarly tuned chain, cranked for maximum, will not give the expected additional 15 dB gain. The overall loop is already operating close to the maximum gain defined by the stopper resistors. A notable measured result is for the maximum boost-and-cut capability of a sweep-mid bell curve to be restricted at the extent of its range where it overlaps across the shelving high-frequency and low-frequency curves.

  A rough rule born from hard experience of squeezing the most EQ from the least electronics is to not allow overlap incursion beyond the point where either curve has ±6 dB EQ effect individually. Overlapping is best achieved from the comfort of another EQ stage, although that too invokes other compromises.
- Hangup 2 is noise. The basic Baxandall, using purely passive frequency-determining components, is a fairly quiet arrangement. With controls at flat, it is theoretically only 6 dB noisier than the unity-gain noise of the amplifier plus additional thermal noise due to network resistances—all in the −100 dBu region. The noise character varies with the controls, as would be expected of an amplifier whose gain is directly manipulated at the frequencies in question—high-frequency boost, more high-frequency noise, and so on.

As soon as active filtering is involved, more noise is unavoidably introduced, often highly colored and consequently much more noticeable. What is worse is that it's present all the time irrespective of control positions. Even with its appropriate control at neutral center, it is quite usual to hear a midsweep swoosh in the noise changing with filter frequency. This is, along with the strange spectral character of the noise emerging from some filters, notably, the integrator-loop variety, a result of unoptimized impedances and dubious stability almost inherent to their design.

25.11.26 Swinging Output Control

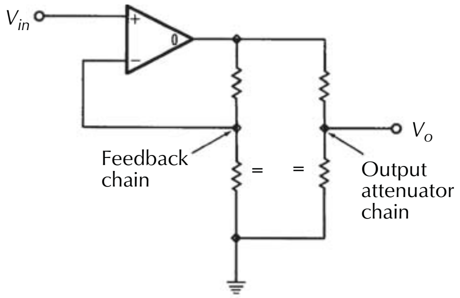
The source impedance versus feedback impedance ratiometric approach of the Baxandall is not the only way of achieving symmetrical boost-and-cut, as stunning as its simplicity and elegance may be. A method of enclosing the controls within the feedback leg of a noninverting amplifier is developed in Fig. 25-69. This has the advantage of leaving the noninverting input of the op-amp free, obviating the need for a preceding low-impedance source or buffer amplifier. Roundabout to this swing is the necessity of a buffer amplifier or quite high destination load impedance since the output is variable in impedance and included within the feedback loop of the op-amp. Serious control law modification, potential phase margin erosion with consequent instability, and certain head room loss are among the penalties for careless termination.

Unity gain in Fig. 25-69A is achieved when the attenuation in the feedback chain equals the output attenuation; the feedback attenuator causes the op-amp to have as much voltage gain as the output attenuator losses. Replacing the two bottom legs of the attenuators with a swinging potentiometer, Fig. 25-69B, provides a boost-and-cut facility; when the pot is swung toward min, the feedback leg is effectively lengthened to ground, and the amplifier gain is reduced somewhat. Meanwhile, the output attenuator is shortened considerably, reducing the output accordingly. At max the reverse occurs. The feedback leg is shortened, increasing the loop gain of the op-amp while the output attenuator is lengthened, losing less of the available output. A small stopper resistor defines the overall gain swing about unity, which would otherwise range from zero to earsplitting, respectively.

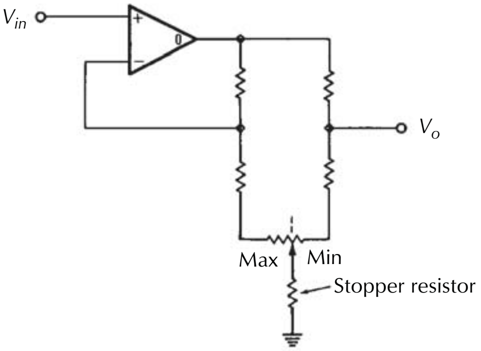
Introducing reactances and complex impedances into the potentiometer ground leg (or legs as in Fig. 25-69C) results again in boost-and-cut control over the frequency bands in which the reactances are lowest, i.e., high frequency for capacitors, low frequency for inductors (real or fake), and so on. This arrangement, which is in a few professional systems and in some Japanese hi-fi, has only one major drawback other than the previously mentioned output-loading considerations. In order to achieve reasonable control dB-per-rotation linearity, the two attenuators (feedback and output) need to be of about 3 dB loss each with the control at center. This implies that the obtainable output voltage is 3 dB below the output swing capability of the op-amp, landing a head room deficit of that amount in the equalizer stage—probably where it is most needed.

25.11.27 Swinging Input Control

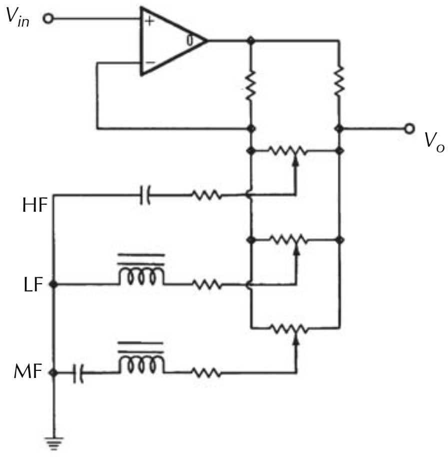
Avoiding the head room headache but utilizing a rather similar technique, the swinging-inputs gain block of Fig. 25-70 is very promising. Here, the feedback attenuator remains unchanged, but the output attenuator is



A. The two chains have equal attenuation for unity gain.



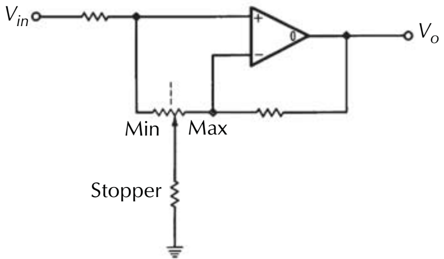
B. Unity at the center, variable gain around unity.



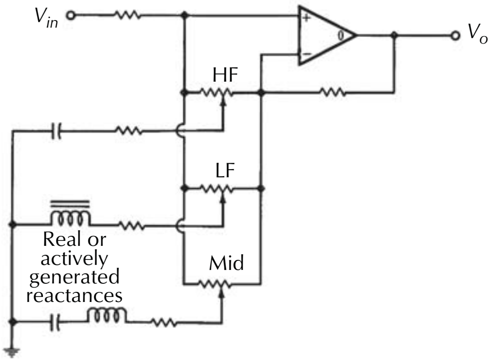
C. Multiband equalizer system approach.

Figure 25-69. Swinging output equalizer.

shifted around to the noninverting input of the op-amp. At minimum, the input attenuation is quite vicious while the feedback leg is long, making the op-amp deliver only a small amount of gain. When the attenuation characteristics are reversed for maximum, the op-amp works at a high loop gain, while the input is only slightly attenuated. Unity is achieved at control center where the input attenuation equals the make-up gain of the amplifier.



A. Unity at center.



B. A three-band equalizer configuration.

Figure 25-70. Swinging input equalizer.

There is a fascinating tradeoff between noise mechanisms in this circuit arrangement. Assuming a maximum of three controls (for fairly basic high-frequency, low-frequency, and midsweep curves) before interaction becomes a major hassle, the amplifier can have between 10 dB and 20 dB of fairly frequency-conscious background gain (i.e., with all controls flat) rendering it at first sight significantly noisier than a Baxandall. However, the impedances around the amplifier are around a decade lower. This considerably reduces thermal noise generation due to resistive elements and op-amp internal mechanisms.

In addition, the noise generated by the active frequency-determining filters is, with the controls neutral, injected equally into the inverting and noninverting inputs of the op-amp. Differential amplifiers being what they are, common-mode signals (such as this equally injected filter noise) get canceled out and do not appear at the output.

Interaction can still intrude, and care is required to prevent excessive frequency band overlap. Center-tapped pots (the tap grounded) eliminate many interactive effects but at the cost of increased invariable background gain (noise) and peculiar, almost intractable, boost-and-cut gain variation linearity versus control rotation.

25.11.28 A Practical EQ

A three-section parametric EQ with additional versatile shelving-type high- and low-frequency controls is detailed in Fig. 25-71. It is designed to be easily shortened to high-frequency, low-frequency, plus a single midband parametric section, for applications that don't demand the full complement of facilities. Each individual section is switchable in or out to allow preset controls. Simple in-and-out comparisons with tie-down resistors maintain the dc conditions of the unused filters to minimize switch clicks. Even a brief look at the circuit reveals a major benefit. The signal path through the EQ is merely via three op-amps, IC2 is an input differential amplifier, and IC3 does duty as the output line amp. In the shortened version this path is reduced to only two op-amps, IC1 and IC3, which serve also as a swinging-input EQ gain block. IC2 and its associated circuitry are unused in this simplified version.

The unusual components around the differential input stage provide unity differential in unbalanced out levels while providing an identical impedance (with respect to ground) on each of the two input legs. Naturally, the more precise the component values, the better the common-mode rejection is likely to be.

25.11.29 The First EQ Stage

IC2 in Fig. 25-71 is the first swinging-input stage. It has two nonfrequency overlapping filters hanging off it, one section covering 25 Hz to 500 Hz, the other covering 1 Hz to 20 kHz. Each filter network creates a complex impedance form against frequency that looks like a series *LC*-tuned circuit to ground. This fake-tuned circuit (formed from two constant-amplitude phase-shift networks in a loop, named the CAPS-variable filter) reach parameters ordinary filters cannot reach.

The center frequency is continuously and smoothly variable over its range using reverse-log potentiometers; *Q* remaining consistent over the entire swing. The *Q* itself is continuously variable between 0.75 and 5 (very broad to fairly sharp, representing bandwidths of 1.5 to 0.2 octaves, respectively). Positive feedback inside the loop, which defines the *Q*, is balanced against negative feedback, which controls minimum filter impedance and, correspondingly, amplitude. Interestingly enough, this circuit relies on the input impedance of the swinging-input stage as part of the negative feedback attenuator. Fortunately, this impedance is reasonably constant irrespective of boost-and-cut control positioning.

In the absence of complementary square-law/reverse square-law dual-gang potentiometers ideally required for the purpose, readily available log/antilog dual-gang pots, retarded a bit to a reasonable approximation, control the positive/negative feedback balance. As a result of this compromise, the filter crest amplitude (maximum effect) varies within ± 1 dB as the *Q* control is swept; in comparison to the dramatic sonic difference from such a *Q* variation, this tends to insignificance. The result of all this, at the output of IC2, is a pair of resonant-type curves of continuously variable place, height, depth, and width.

25.11.30 Second EQ/Line Amp

A reasonably hefty pair of transistors is hung on the end of IC3 to provide a respectable line-drive capability, in addition to the use of the amplifier as a swinging-input EQ section. There is enough open-loop gain in the combination of the op-amp and transistors (over a much greater bandwidth than mere audio) to cope with 15 dB of EQ boost and output-stage nonlinearities.

Differing from the last EQ stage, this one only has a single midfrequency bell-curve creator, operating over a range of 300 Hz to 3 kHz, together with high- and low-frequency range impedance generators.

25.11.31 Low-Frequency Control

Gyrating inductance to create a conventional low-frequency shelving response (variable in turnover frequency by a 220 kΩ antilog pot) is achieved around IC11. A fairly large (2.2 µF) series capacitor forming a resonance is switchable in and out. The value of the capacitor is carefully calculated to work with the circuit impedances to provide an extreme low-frequency response that falls back to unity gain below the resultant resonant frequency without compromising the higher frequency edge of the curve. The *Q* of this arrangement reduces proportionally to increasing frequency. Typical resultant response curves, Fig. 25-72, show just what all this means, demonstrating an extraordinarily useful bottom-end control.

25.11.32 High-Frequency Control

Unusual is one way to describe the high-frequency impedance generator and its EQ effect. It is essentially a supercapacitor, or capacitative capacitor. In other words, it's a circuit that, when in conjunction with a resistor, causes a second-order response as would normally be expected of an inductor and capacitor combination—a slope of 12 dB/octave as opposed to a single-order effect of 6 dB/octave. Fig. 25-73 shows what it does as an EQ element.

The response is hinged about 1 kHz. The control varies the frequency (between 5 kHz and 20 kHz) at which the gain reaches maximum (or minimum if the boost-and-cut control is cut). The slope between 1 kHz and the chosen maximum frequency is virtually a straight line representing a nearly constant dB/octave characteristic, with a nearly flat-top shelving characteristic.

In electronic terms, this is achieved by progressively degenerating the supercapacitor until it's no longer super—i.e., it eventually ends up looking like a simple, single capacitor.

25.12 Dynamics

Dynamics processing is becoming as common in today's signal paths as equalization. Many commercial mixing consoles carry dynamics as standard on a per-channel basis. Formerly, the occasional necessity of hanging in external dynamics processing was the major justification for channel insert (or patch) points. These patch points were typically either before or after (pre- or post-) equalizer in the signal chain as shown in Fig. 25-74.

Preequalization actually amounts to postinput stage; the purpose of an insert point here is to control unruly input signals that may otherwise clip later in the chain—in particular within the EQ section where serious amounts of gain at some frequencies may be instituted. Characteristically this pick-off point is *after* any high-pass filtering, so that any low frequency rumble that may cause false triggering within the dynamics may be removed. On the other hand, the post-EQ insert point allows control over the entire channel immediately prior to the channel fader and is commonly used for automatic gain controlling.

Dynamics is automatic control of signal level by an amount determined by the characteristics of the signal itself. In a linear 1:1 circuit, what goes in comes out untouched and unhindered. For example, if we have a circuit that automatically senses the input signal and uses that measurement to control the output signal, if the input signal rises in level by 6 dB, the output signal is controlled to rise only 3 dB. The output signal has been compressed by a ratio of 2:1 with respect to the input signal.

There are four basic types of dynamic signal processing:

1. Limiting.
2. Gating.
3. Compression.
4. Expansion.

It is arguable that limiting is a special case of compression and that gating is similarly a special case of expansion, effectively reducing the number of basic groups to two. Although in practice the means of achieving these pairs of effects are indeed similar, true definitive compression is a long way from limiting, as gating is from expansion. The discussion about time constants etc. within the following Limiting section is directly applicable to each of the other sorts of dynamics.

Fig. 25-75 shows a now customary style of input-output signal level plot of a compound (more than one dynamics type active) dynamics section, with typical slopes for each of limiting, compression, and expansion. (Measured plots of actual dynamics sections can be seen also in the discussion of digital dynamics later.) This style of dynamics display is common on programmable consoles; here it can be a handy form of visualization or representation as the various types are discussed. Linear (i.e., no processing) is represented by the dotted line, showing equal output for input. Unusually (!) a section of linear remains in this compound curve, displaced upward by 15 dB; this is a normal occurrence when automatic gain reduction such as compression and/or limiting is used. In order to make up for the gain reduction above the threshold, some gain is applied to compensate and bring the output signal back up to usable levels. This is often called makeup, or buildout gain.

25.12.1 Limiting

This is the conceptually easiest and most commonly applied form of dynamics processing. Nearly any audio heard outside of a recording studio has been passed through a limiter at some point in the chain—TV audio, radio, even (and especially) CDs. This very pervasiveness has had the rather odd unforeseen effect that culturally we have gone beyond acceptance and become dependent on the sound of excessive limiting and heavy compression; even to educated ears a new CD can sound wrong or wimpy in comparison to what had been heard on the radio, where typically murderous additional processing is applied. Record shops actually do get complaints like that. The race to be louder than loud has unfortunately spilled over into the record production arena back from radio where it had previously raged

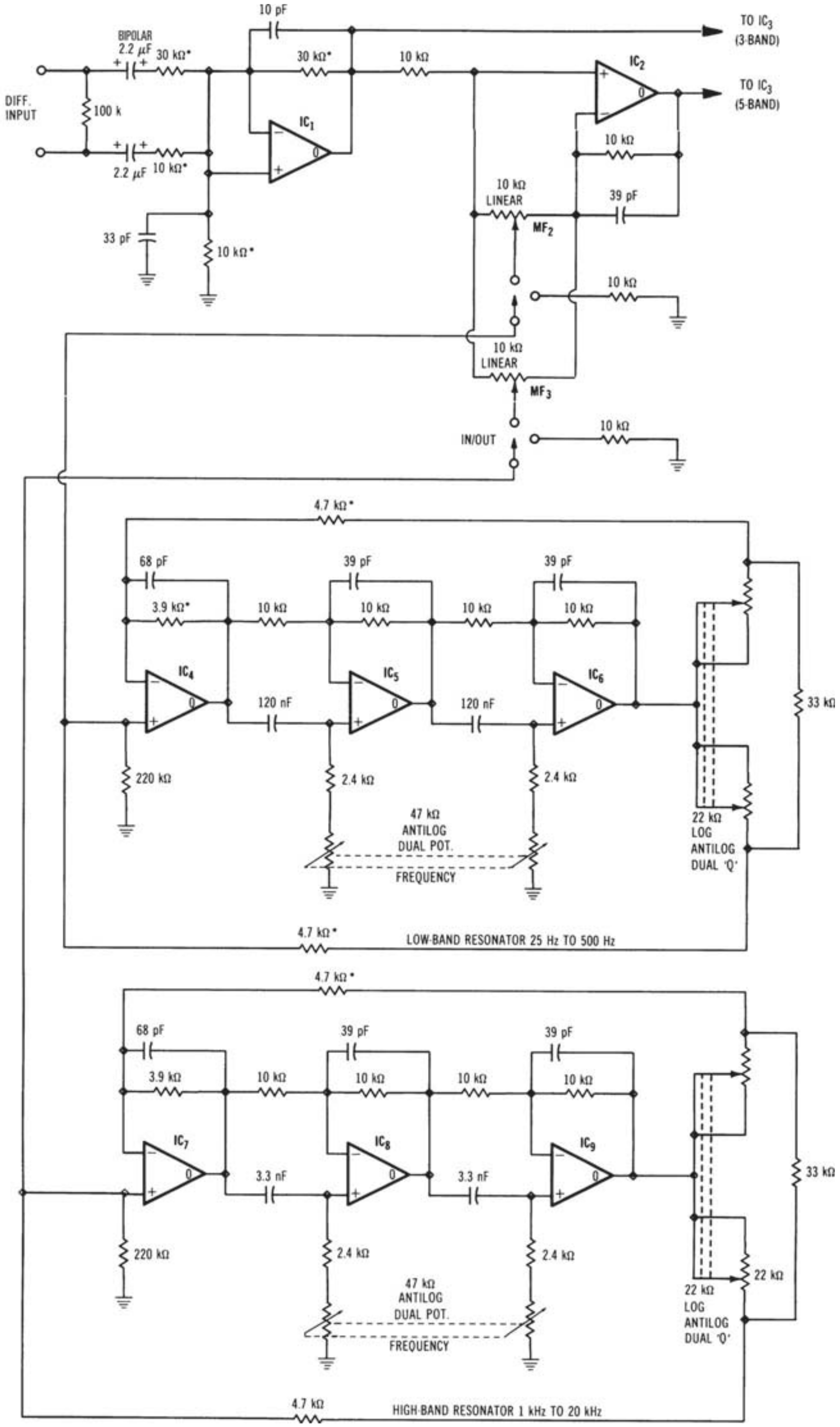

10 pF
TO IC_3 (3-BAND)
TO IC_3 (5-BAND)
BIPOLAR 2.2 μF
30 kΩ*
30 kΩ*
10 kΩ
IC_2
DIFF. INPUT
100 k
IC_1
2.2 μF
10 kΩ*
10 kΩ
10 kΩ LINEAR
39 pF
MF_2
33 pF
10 kΩ*
10 kΩ
10 kΩ LINEAR
MF_3
IN/OUT
10 kΩ
4.7 kΩ*
68 pF
39 pF
39 pF
3.9 kΩ*
10 kΩ
10 kΩ
10 kΩ
10 kΩ
IC_4
IC_5
IC_6
120 nF
120 nF
33 kΩ
220 kΩ
2.4 kΩ
2.4 kΩ
47 kΩ ANTILOG DUAL POT.
FREQUENCY
22 kΩ LOG ANTILOG DUAL 'Q'
4.7 kΩ*
LOW-BAND RESONATOR 25 Hz TO 500 Hz
4.7 kΩ*
68 pF
39 pF
39 pF
3.9 kΩ
10 kΩ
10 kΩ
10 kΩ
10 kΩ
IC_7
IC_8
IC_9
3.3 nF
3.3 nF
33 kΩ
220 kΩ
2.4 kΩ
2.4 kΩ
22 kΩ
47 kΩ ANTILOG DUAL POT.
FREQUENCY
22 kΩ LOG ANTILOG DUAL 'Q'
4.7 kΩ
HIGH-BAND RESONATOR 1 kHz TO 20 kHz

Figure 25-71. Five-band equalizer circuit diagram.

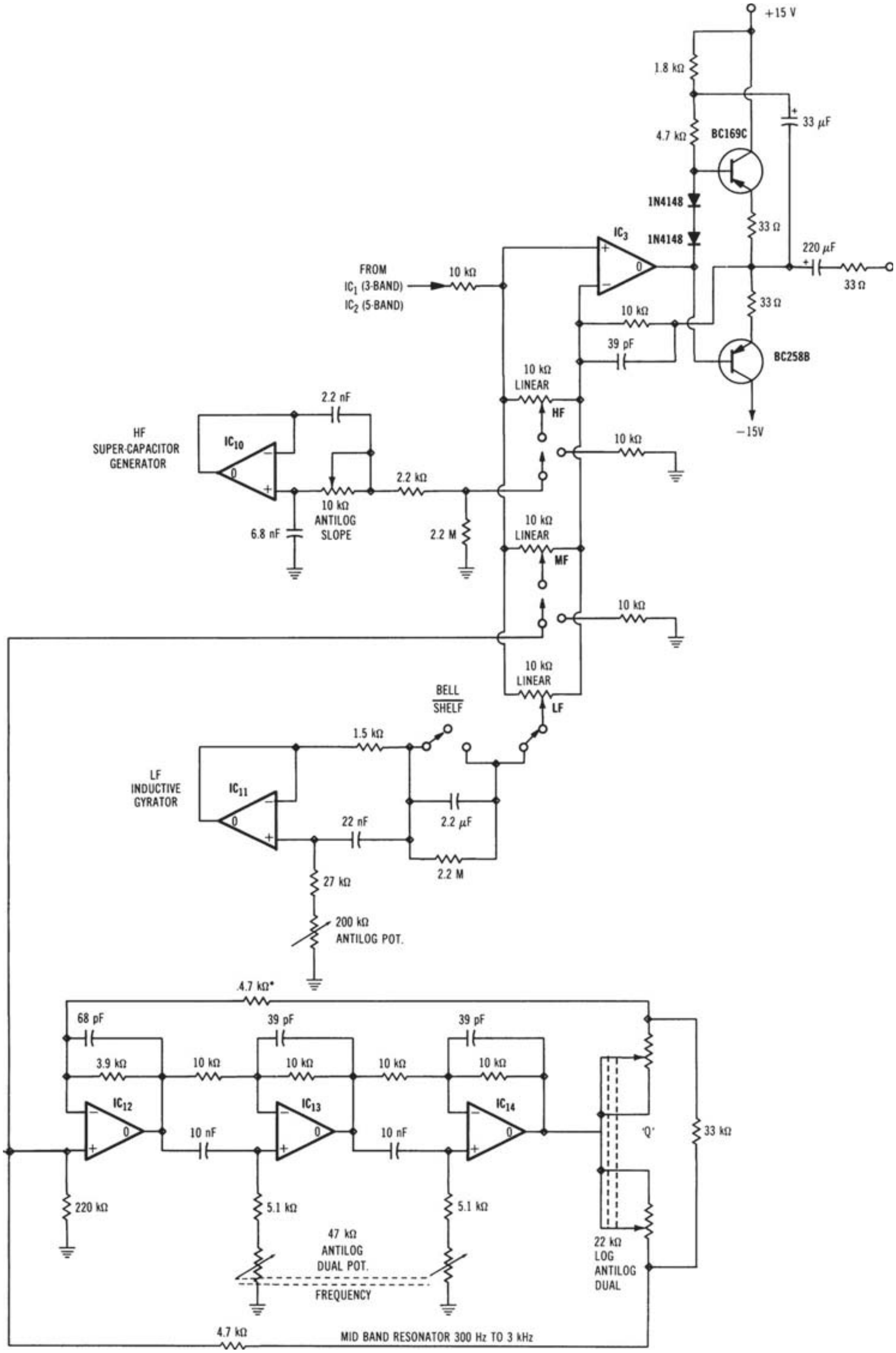

+15 V
1.8 kΩ
33 μF
4.7 kΩ
BC169C
1N4148
1N4148
33 Ω
IC3
220 μF
33 Ω
FROM
IC1 (3-BAND)
IC2 (5-BAND)
10 kΩ
0
33 Ω
10 kΩ
39 pF
BC258B
−15V
10 kΩ
LINEAR
HF
2.2 nF
HF
SUPER-CAPACITOR
GENERATOR
IC10
0
10 kΩ
2.2 kΩ
10 kΩ
ANTILOG
SLOPE
6.8 nF
2.2 M
10 kΩ
LINEAR
MF
10 kΩ
10 kΩ
LINEAR
LF
BELL
SHELF
1.5 kΩ
LF
INDUCTIVE
GYRATOR
IC11
0
22 nF
2.2 μF
2.2 M
27 kΩ
200 kΩ
ANTILOG POT.
4.7 kΩ*
68 pF
39 pF
39 pF
3.9 kΩ
10 kΩ
10 kΩ
10 kΩ
10 kΩ
IC12
IC13
IC14
0
0
0
10 nF
10 nF
'Q'
33 kΩ
220 kΩ
5.1 kΩ
5.1 kΩ
47 kΩ
ANTILOG
DUAL POT.
FREQUENCY
22 kΩ
LOG
ANTILOG
DUAL
4.7 kΩ
MID BAND RESONATOR 300 Hz TO 3 kHz

Figure 25-71. Five-band equalizer circuit diagram. Continued.

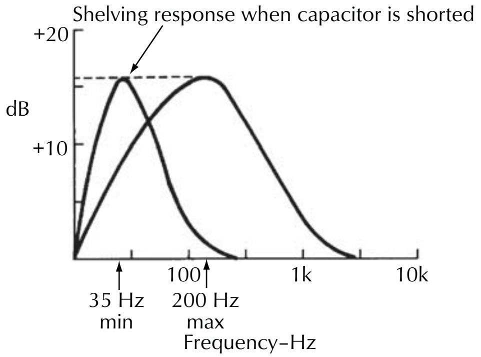


Figure 25-72. Frequency response of the low-frequency section of Fig. 25-71 (control at maximum gain).

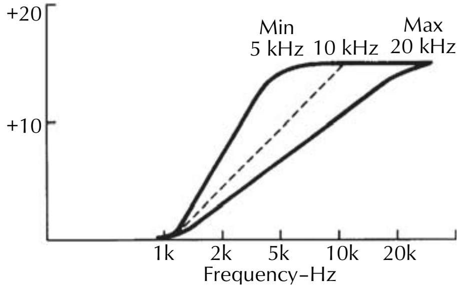


Figure 25-73. Characteristics of the high-frequency section of Fig. 25-71 (boost-cut control at maximum boost).

alone sparked by the AM radio loudness wars of the ’60s and ’70s; the deleterious effect will have future musical historians scratching their heads.

Every recording and transmission medium has definite head room limitations, a maximum level beyond which the signal just plain overloads, distorts, or becomes a serious liability. AM radio overmodulation not only sounds horrid but causes interfering splatter up and down the dial, FM transmitter overdeviation causes adjacent channel interference and runs the risk of distortion in receiver discriminators; disk cutters can produce grooves that run into each other (long after they become unplayable by any normal pickup), PA loudspeakers fry, tape saturates, distorts, and screams, and best of all anything digital just plain runs out of bits and cracks up. The answer? A device that senses when enough is nearly enough and automatically reduces the source level such that a proscribed output level is not transgressed. This is a limiter.

Fig. 25-76A shows the all-time basic limiter—a pair of back-to-back diodes. These clamp the input signal from the source resistor to within their nonconductive range; beyond 700 mV of either polarity one or other of the diodes conducts, sawing off any excessive signal. Brutal, but effective. The downside is gross distortion—serious waveform modification is going on, profligate audible distortion products are generated. Fig. 25-76B shows the same idea with germanium diodes. These tend to have a lower turn-on voltage (200–300 mV) but a gentler knee, with the effect shown. This sounds considerably less harsh—fewer high-order distortion products are being created. In situations where ultimate signal quality is not necessary, but increased signal density (translated: loud) is required, these clipping circuits work like a charm; communication circuits often use this technique to saw the top 10 dB or so peaks off speech and thus gain a nearly corresponding degree of increased apparent loudness. The trick is to filter away or contain and control the resulting distortion products such that they become less agonizing, while retaining the high signal density clipping affords. Such is meat for a whole other saga.

The ideal circuit is one that knows its input is going over the top and can reduce its gain such that the signal is left relatively undistorted but as loud as it can be within the given constraints. Fig. 25-76C is a block diagram of such a device. The side chain circuit is a tripwire in this instance; if the amplifier output exceeds a stipulated level (just below what the destination can handle) the side chain develops a control signal that

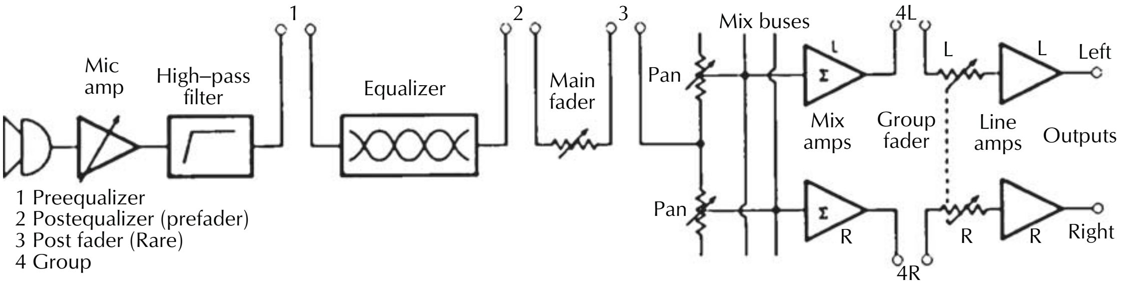


Figure 25-74. Typical insert points for dynamic processing.

tells the input attenuator to drop the input signal sufficiently. The whole circuit operates in check—the bigger the input signal, the bigger the potential overload, the bigger the control signal, the more the attenuation. Below the tripwire—the threshold—the whole circuit behaves the same as an ordinary straight amplifier. Fig. 25-76D is about as simple as a decently performing limiter can get and was first noted in the original mid-sixties Philips cassette recorder.

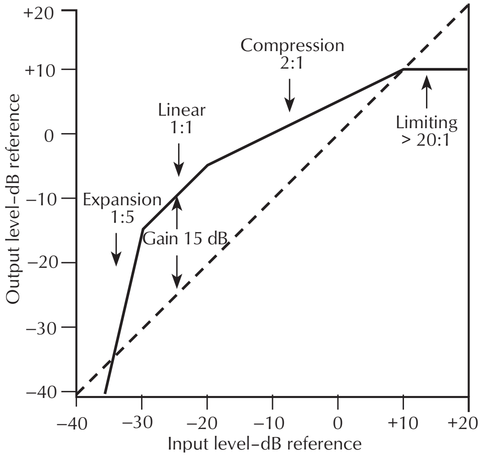
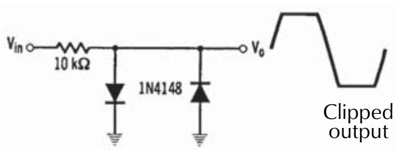
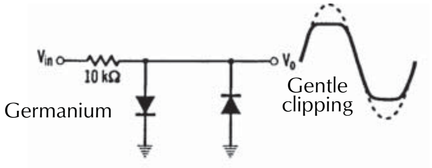


Figure 25-75. Dynamics Input/Output Plot

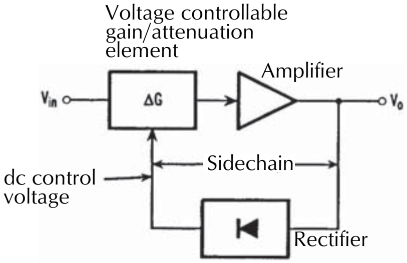
The LM386 is a small power op-amp commonly used to drive headphones or small loudspeakers but works well just as an ordinary amplifier, in this case at some 30 dB gain. It is used here for its power output stage, which is hefty enough it can ignore the diode rectifier and side chain loading effects. This diode conducts when the positive-going output signal exceeds about 700 mV and charges a reservoir capacitor, C_r. This is buffered by an emitter follower (TR_f) feeding the gain controlling transistor TR_g. When the voltage on the base of TR_f is sufficient to force conduction through the two base-emitter junctions, TR_g turns on, causing an increasingly low-impedance path to ground at the input to the amplifier. It forms a potentiometer, with the source resistor R_s attenuating the input signal to the level at which the rectifier and two transistors are just conducting. In this circuit that amounts to a positive-going output signal of about 2 V (the added voltage drops of the rectifier and the two transistor base-emitter junctions). Simplicity has its drawbacks and in this instance they are noise and distortion. Although the distortion is in a different league from a diode clipper, transistors are not ideal VCAs and are somewhat nonlinear in this application. If, however, the signal


V_{in}
10 kΩ
1N4148
V_0
Clipped
output

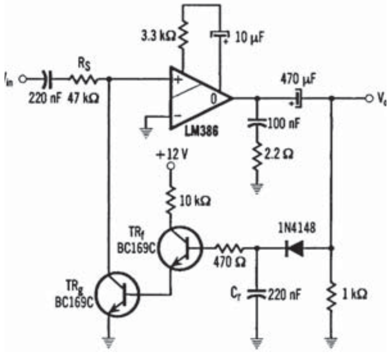
A. Diode clipping.


V_{in}
10 kΩ
Germanium
V_0
Gentle
clipping

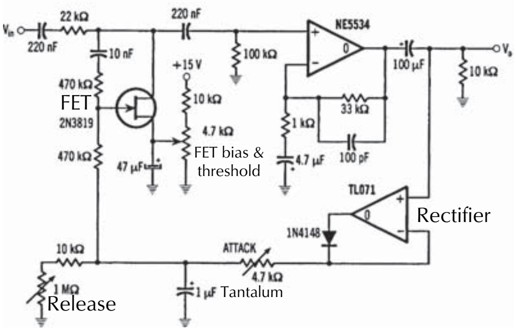
B. Soft diode clipping.


Voltage controllable
gain/attenuation
element
Amplifier
V_{in}
ΔG
V_0
Sidechain
dc control
voltage
Rectifier

C. Block diagram of a feedback limiter.


3.3 kΩ
10 µF
R_S
V_{in}
220 nF
47 kΩ
0
LM386
470 µF
V_o
100 nF
2.2 Ω
+12 V
10 kΩ
TR_f
BC169C
TR_g
BC169C
1N4148
470 Ω
C_r
220 nF
1 kΩ

D. Simple feedback limiter.


V_{in}
220 nF
22 kΩ
10 nF
220 nF
100 kΩ
+15 V
NE5534
0
100 µF
V_o
10 kΩ
470 kΩ
FET
2N3819
10 kΩ
33 kΩ
1 kΩ
4.7 kΩ
FET bias &
threshold
470 kΩ
47 µF
4.7 µF
100 pF
TL071
Rectifier
1N4148
10 kΩ
ATTACK
4.7 kΩ
1 MΩ
Release
1 µF Tantalum

E. An FET-based limiter.

Figure 25-76. Simple limiters.

across them is kept low (in this instance, –30 dBu, the lower the better) it can be quite acceptable for a lot of applications. Keeping the signal low necessitates following gain to bring the signal back up again. That means amplifier noise.

There is a wide variety of possible voltage-control elements and Fig. 25-76E shows a smoother version of the transistor limiter using a JFET. The principle is much the same, only the side chain and VCA circuitry has developed somewhat. FETs have such spread characteristics that a preset is necessary to set their bias points. In normal operation the FET needs to be biased just nonconducting; that is, nonattenuating. This necessary adjustment also provides a means of varying the output level at which the limiter starts to bite, and incidentally some control over the ratio of the gain reduction. (The greater the bias, the more control voltage signal is needed to be generated before the FET turns on and starts attenuating, where it does so in a tightly controlled manner. A low bias results in a lower, or indeed no, threshold and a far gentler gain-control ratio. Carefully trading these two— a high threshold for a hard ratio or low and "smushy" threshold and gentle FET turn-on—against overall gain formed the basis of FET-based compressors such as the famed Audio and Design F760 and the UREI 1176LN.)

FETs have very high gate impedances, precluding the need for a follower. The control voltage is summed at the device gate with a sample of the input signal. Automodulation is an effect of FETs where the source-drain resistance (the resistance we're depending on as part of the input attenuator) varies with signal voltage across it. This is attacked in two ways: first by keeping the signal across the FET low, as in the transistor limiter, and second by supplying the gate with some anti-wobble signal that does a fairly good job of forcing the source-drain path to wobble against and largely cancel the automodulation effect.

25.12.1.1 Side-chain Time Constants

Between the rectifier and the FET in Fig. 25-76E is a simple resistor-capacitor network that determines how the side chain works and its effect on the automatic gain reduction. This is in contrast to the transistor circuit where the reservoir capacitor discharges through the transistors at one end and is charged rapidly through the diode from the other. Here we can adjust the rate at which the capacitor charges and at which it discharges. The implications of these on how the circuit behaves and sounds are crucial.

But why have time constants at all? If the idea is to provide protection for overloads, why bother with how they're handled? Well that's all diode clippers do, Figs. 25-76A and B. They have zero attack time, which means there is no delay or run-up to them when dealing with an overload. Similarly, they have a zero release time, meaning that once the overload is dealt with it's instantly business as usual. The trouble is, they sound horrible, as would either the transistor or the FET limiter with infinitely short parameters.

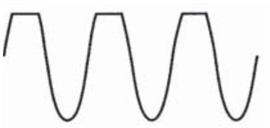
25.12.1.2 Attack Time

Fig. 25-77 shows the first few cycles of a train of sine waves that are in excess of the limiter threshold and the effect as the limiter tries to reduce the output to the prescribed level. Fig. 25-77A shows a zero attack time and not unexpectedly looks very sawn off. Lengthening the attack time somewhat, Fig. 25-77B, leaves a recognizable but mutilated crest, while longer still is even less bent, Fig. 25-77C. Unfortunately, the character of distortion products generated by this effect are very audible; they are loud (since it is a loud signal that is subject to control) and of high order and unlikely to be masked by the fundamental signal. Even at this stage it is clear that a longer attack time takes less toll of the input signal integrity; the less a waveform is modified, the better it will sound. Expanding the time scale to many cycles shows how lengthening attack time looks. A long attack time, Fig. 25-77F, gradually reduces the limiter gain until the signal is completely under control while imposing less immediate distortion on it.

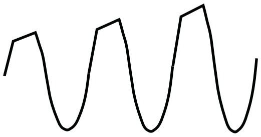
The tradeoff is apparent. An attack time long enough to not mangle the program material also permits excess output level for as long as the circuit takes to bring the gain down sufficiently. Balancing this overshoot against leading-edge distortion due to short attack times is a subjective compromise.

Naturally the lower the frequency, the longer the time period between cycle crests and the greater distorting effect of attack times. An adequate attack-time for high frequencies can easily be still way too short for low frequencies, while an adequate attack time for bass is unnecessarily long for highs. That's life.

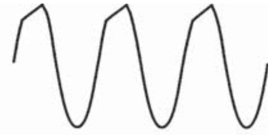
It is normal in a high-quality studio dynamics section to use a full-wave rectifier for the side chain (as opposed to the half-wave shown in these two examples). This gives twice as many opportunities per cycle to sense and adjust the gain (one on the positive-going peak, one on the negative) in addition to allowing for the fact that few real-world signals are symmetrical;



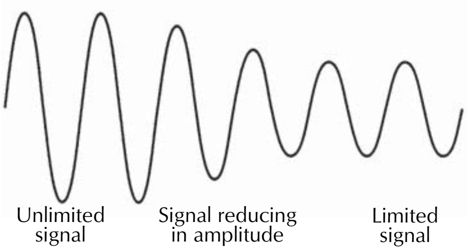
A. Zero attack time.



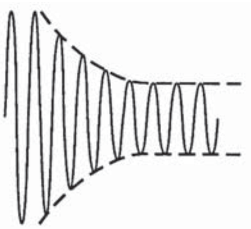
B. Medium attack time.



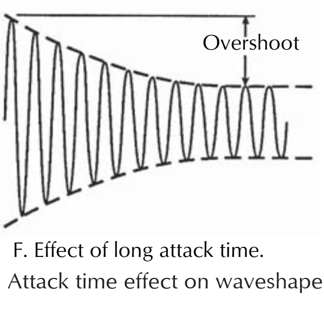
C. Long attack time.



D. Effect of zero attack time.



E. Effect of medium attack time.



F. Effect of long attack time.

Figure 25-77. Attack time effect on waveshape.

either the positive or negative peaks are more, sometimes greatly, pronounced.

Attack times are measured and quoted either in microseconds or milliseconds (i.e., the time constant of the reservoir capacitor versus charging resistor, which also corresponds approximately to the time a transient takes to be controlled) or alternatively in dB/ms which is the rate at which the attenuation changes.

25.12.1.3 Release Time

The purpose of a release time constant is manifold, but in the case of a peak limiter its value is primarily to minimize distortion, much as the attack time. If the incoming sine wave train is above the threshold and the limiter is trying to contain it, and if the release time is short, the gain will tend to recover between each individual crest of the sine wave. In effect, this is the reverse of the attack-time distortion. Although there is the brutality of a fresh set of attack-time-related distortion on each crest, it must not be forgotten that as the attenuation releases there is a less traumatic but nevertheless real change in shape of the rest of the waveform as its amplitude changes within a cycle.

Release times are normally maintained much longer than attack times. With the exception of true transients, which spike up once and then go away for an indeterminate period of time, if not forever, most sounds tend to stay around for a while—at least for a few cycles. It can be reasonably assumed that, once a signal has hit threshold, more of it will follow; given that, there is little point in letting the attenuation drop back just to be reasserted milliseconds later. The release time is a crude memory of the size of the signal the section is having to deal with at a given moment and by keeping the amount of attenuation relatively stable gives less work to and less damage to wreak for the attacking charge ramp-up. A longer release time constant gives the attack circuitry less to do except at the onset of material over the threshold.

There is always the danger with long release times (if chosen to minimize distortion) in that should a large transient come along, the limiter will do its job and promptly reduce the gain to prevent excess output level. Fine, but the long release time keeps the attenuation invoked sitting around for a long time, compressing following program material and a large amount of following information can be lost until the gain claws its way back up to normal level, Fig. 25-78.

The subjective compromise between long release times (for distortion's sake) and rapid recovery depends largely on the program material. More so with release than with attack; too short a release time can really tear up bass frequencies. Distortions due to attack and release time constants—transient, intermodulation, and harmonic—owe themselves to the fact that gain is

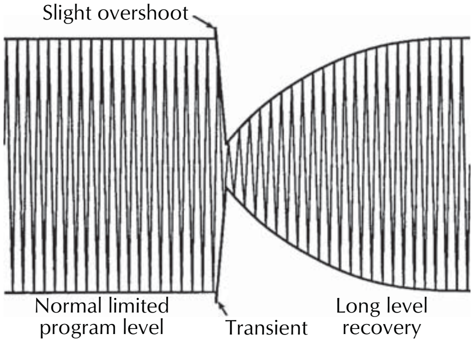


Figure 25-78. Effect of long release times on the output signal when subject to a large transient.

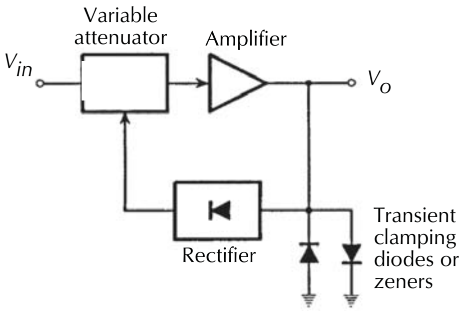
changing and rapidly. They are independent of the kind of device doing the attenuation, whether it is a humble transistor or an expensive VCA and they are just as obnoxious. By and large these dynamically induced distortions subjectively far outweigh the steady-state distortion characteristics of the devices. These only become important if the circuit is to sit in a signal path with little or no dynamics processing taking place. Once things start moving one is as good or bad as the other and the subjective quality of a unit is determined by how well the various timings tailor around the program material or how well it does so automatically.

Release settings are generally quoted in milliseconds or seconds, and sometimes as a decay rate such as dB/ms or dB/s. In short, a dynamics processor lives or dies by its side chain.

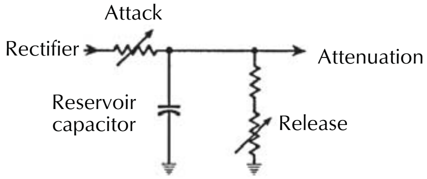
25.12.1.4 Compound Release Time Constants

The hole-punching problem of a big transient hitting a limiter with a long release time constant can be attacked in a couple of ways. If subtlety is none too great a criterion—for instance, in an AM transmitter limiter where *There Shall Be No Overmodulation*—it is always possible to put a variation of the back-to-back diode limiter on the output of a timed one, Fig. 25-79A. Set to clip immediately above the normal operating output level of the feedback limiter, it not only cast-iron stops excessive output signal swing but also prevents the transient from entering the side chain and digging too big a hole in the following audio.

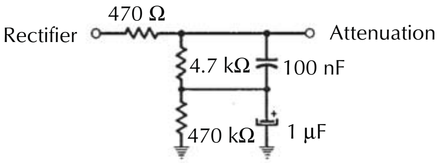
The circuit in Fig. 25-79C is used extensively. Instead of a single reservoir capacitor in the side chain with one attack defining resistor and one release defining resistor, Fig. 25-79B, a compound circuit can be arranged, Fig. 25-79C. A small value resistor and capacitor form additional, shorter attack and release time constants working in conjunction with a slower set. The extended attack and release times follow the general loudness envelope of the program material while the shorter ones, riding on the top take care of any short-term discrepancies and transients, generally to the tune of the top 5 dB or so of processing. If a general-purpose, hands-off, no-tweaks limiter is needed, this arrangement with carefully chosen values can work very nicely and is the basis of some commercial outboard limiters in the auto mode.



A. Feedback limiter with output clamping.



B. Simple attack/release sidechain.



C. Compound time-constant side chain arrangement.

Figure 25-79. A feedback limiter with a compound time constant.

25.12.1.5 Multiband Processing

A complex but very effective side step to the inapt time constant with frequency problem is shown diagrammatically in Fig. 25-80. Here the input signal is split a number of ways by frequency (in this case three, into bass, mids, and highs). Each band passes into a limiter with side chain time constants optimized for that band of frequencies; this allows the highs to be optimally treated with short time constants without compromising the other sections, and so on. The bands are recombined and passed through an envelope limiter, which only needs to catch a limited dynamic range and so has little effect on the path and the overall sound. Two bands is enough to remove the pumping effect of the usually energy-intensive bass-modulating higher frequencies; three bands allow for better time-constants for the all-important mid range to be established without compromising the high and low frequencies; more bands, say five or six, allow considerable program density (translation: loudness) to be built up while retaining musicality.

This is a very common technique in radio broadcasting airchain processing, and allows for better or worse far deeper processing than possible with broadband units and totally avoids side chain pumping effects, where typically heavy bass modulates the mids and highs content. It is usual for there to be a number of multiband stages, preceded by broadband AGC and succeeded by broadband limiting/clipping. The first multiband section (five bands is common) being compression and perhaps multiband AGC, feeds a second section of multiband limiting (31 bands of limiting is not unknown!). Needless to say, such devices can be quite an entertainment to set up; indeed, a whole subindustry of processor witchcraft has evolved in radio.

Adjusted well, these units can sound startlingly good (and loud). In corollary, they are far easier to make sound truly dire. Unfortunately, the multiband technique, either using discrete units such as air-chain processors or virtually as software plug-ins to audio workstations and digital consoles, has found its way into music production. The results have rarely been beneficial.

25.12.1.6 Active Release Time Constants

Passively discharging the side-chain reservoir capacitor with a resistor is not necessarily the best way of going about things. Looking at Fig. 25-81A shows that the initial discharge rate is considerably faster than that farther along in time. With a gain control element (i.e., a Voltage Controlled Amplifier/Attenuator—VCA) having a linear control voltage to dB attenuation characteristic, the gain reduction on release would die away very quickly initially and steadily bottom out. This is bad news since a longer than necessary release time constant would need to be applied to preserve adequate low-frequency distortion. If the reservoir capacitor is discharged linearly as in Fig. 25-81B by a constant-current source instead of by a straight resistor (example in Fig. 25-81C), a tidy linear dB attenuation release versus time characteristic ensues; less release time need be wound in for similar LF distortion.

This can be taken a step further. Some gain-control elements with logarithmic (transistors) or square-law (FET) control voltage characteristics, for example, can be made to work with a passive release system to give a pretty good approximation of linear dB/time release. Adding constant-current discharge to one of these circuits gives a slow discharge initially (i.e., good low frequency distortion) with a more rapid tailoff, Fig. 25-81C, which removes unnecessary gain reduction quicker than any other arrangement yet. On program material this works very well, also serving to reduce pumping and suck outs from transients.

25.12.1.7 Hold or Hang Time

Given active discharge with a constant current source, it is always possible to turn the discharge path off. This has the effect of freezing the attenuation at the instant

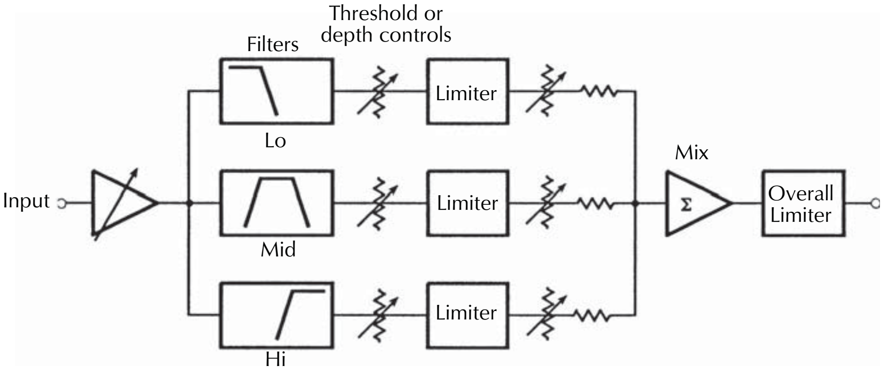


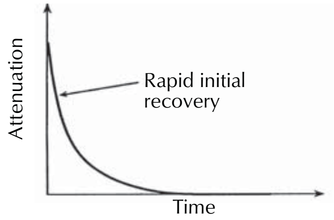
Figure 25-80. Multiband signal processing.

the discharge is removed; if there is no discharge path, the reservoir can't discharge, the control voltage remains static, as, consequently, does the attenuation. Recent refinements to dynamics include this feature; if the side chain is attacking in response to an increasing signal over the threshold, the constant-current discharge is turned off automatically. It remains turned off for a preset amount of time after the attacking has ceased and when the circuit would ordinarily be releasing. Instead, the attenuation remains static for that preset time after which the discharge is reinstituted and normal release decay occurs, Fig. 25-81D. The advantages are straightforward; there is no release time-constant-related distortion in the period of time the attenuation is frozen and the subsequent release can be independently set to return gain as quickly or as slowly as desired. Tailoring the attack, hold-and-release times around a given program source can render processing virtually transparent in many cases. *Hold* or *hang* time is quoted usually as a direct time in milliseconds or seconds.

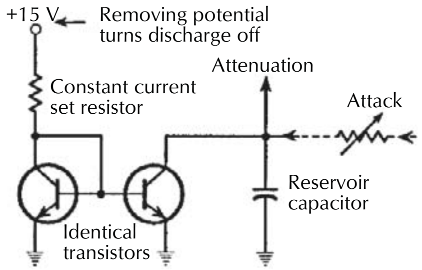
25.12.1.8 Limiting and Compression for Effect

The principal creative purpose for limiting or compression (as opposed to the precautionary and damage-control functions outlined earlier) is to make things loud. Suitable parameters can also imbue low-frequency chunkiness or weight on the sound. Given a certain maximum head room level in a transmission or recording medium, it is often desirable to increase the program density or reduce the dynamic range. A case in point (again) is broadcasting. Their legitimate purpose for compression is to render audible portions of program that may otherwise be too quiet and buried in ambient noise, interference, or static at the receiver end. Automobiles have a notoriously small dynamic window between receiver output capability and cabin noise. A compressor (or usually air chain processor) is set to give sufficiently high output for a quiet program and to automatically reduce the gain when it gets louder.

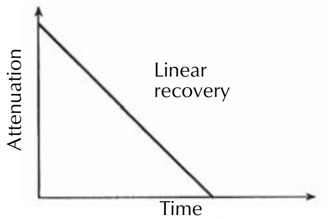
An interesting subjective side effect of this lies in psychoacoustics; if the ear hears something it knows to be quiet ordinarily at a certain volume, then a sound that it knows to be louder still seems louder even though a limiter may be compressing the two signals to the same level. A classic example is with reverberation tails—these are one of many means by which we subconsciously gauge relative loudness; if the original sound spawning the reverb tail is compressed to be closer to the level of the tail, the whole overall sound seems louder.



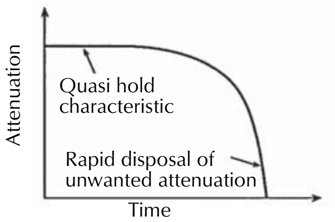
A. Attenuation versus time for simple capacitive discharge (with VCA).


+15 V
Removing potential turns discharge off
Attenuation
Constant current set resistor
Attack
Reservoir capacitor
Identical transistors

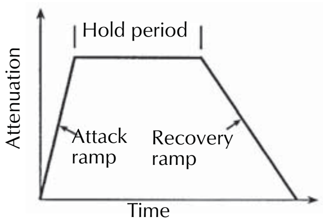
B. Constant current discharges using current mirror.



C. Attenuation versus time for constant current discharge (with VCA).



D. Attenuation versus time constant; current discharge with nonlinear attenuation/control voltage characteristics.



E. Attenuation versus time showing deliberately introduced hold or hang time prior to release.

Figure 25-81. Active release side-chains.

Normal program material consists of quite high transients and peaks above the mean average level. If these peaks are removed (brutally by a clipper or more subtly by a timed processor with short time constants), the average transmitted level can be increased correspondingly. The shorter the time constants, the more apparent loudness can be squeezed out. Usually, though, this is at the expense of quality.

Here belies the principal reason broadcasters like dynamics processing—the louder they seem on the air, the more listeners are attracted to the station. To this end excruciating amounts of gain reduction are common on-air in radio. It's a well-known effect that something that sounds louder—even marginally—is perceived as sounding better, at least in the short term.

It is also a strong reason compression is prevalent in individual recording channels within a console; each sound can be made not only more controlled in level, which helps balancing, but also denser and more solid sounding. The downside is that it's so easy to squash vitality out of a sound, trading liveliness and depth for something more up front but ultimately less interesting.

25.12.2 Gating

Gating is to a degree the inverse of limiting. It is the removal of an output signal unless it is of a sufficient strength; in other words, if the input signal is above a threshold level it is permitted to pass, but if it falls below the threshold it is attenuated.

Its purpose is usually to remove or reduce in level a signal when it is no longer usefully contributing to a mix, remove noise in between wanted sections of program and to generally act as an automatic mute. A true gate totally removes the undesired signal but in practice—for noise reduction in particular—a lesser amount of attenuation is invoked; this is set by a control and indicator called *depth* or maybe just attenuation. Gentle amounts of depth make the operation of a gate far less obvious together with the benefit that there is less intermodulation distortion if the gain is asked to change through less of a range.

The gate attack or wake-up time is generally adjustable and determines how quickly the gate opens in response to a signal tripping the threshold. It is usually set very fast, though, such that none of the leading edge of the signal is missed. The hold time (if there is one available, usually) determines how long the gate remains open after the signal drops below the threshold and the release or decay time sets how quickly the attenuation returns. That these are a direct parallel to their behavior in the limiter is no accident; nearly all dynamics processing sections carry these controls. It only needs to be remembered with a gate that the attack time has to do with how quickly attenuation is removed rather than applied as is the case with a limiter.

The ranges of time-constant values are typically similar to those for a limiter, but the threshold range can extend from 0 dBu (or even as high as +20 dBu in some cases) down to about −40 dBu or below. The higher thresholds are mostly for key triggering while the low extremes are for noise reduction. Automuting settings are somewhat critical, needing to be above the general background level yet below the desired signal's typical level; there is always some tuning to be done, but figures of −10 dBu to −20 dBu are typical. Depth can be adjustable between 0 dB and 40 dB attenuation (some manufacturers optimistically state infinity).

Practical uses include automatic microphone muting (backup singers), spill removal (e.g., a snare drum microphone is usually gated so that when the snare isn't actually being hit, the microphone isn't picking up the rest of the kit), and noise reduction (just enough gating applied to tape track returns to subdue tape hiss or air conditioning rumble). In all cases the parameters are set up to be as unobtrusive as possible. These vary from lightning fast attack and decay on a snare drum to fairly leisurely ramps in noise reduction.

In addition to the hold, or hang, time which prevents the gate from chattering on a marginal signal, an additional tool to prevent falsing is hysteresis between the signal level necessary to open the gate (open threshold) and that below which the gate considers the signal to have gone away (close threshold); this hysteresis (a few dB) is generally concealed from the operator.

25.12.2.1 Gating Feed-Forward Side Chain

Naturally, a gate cannot possibly operate with its side chain taken from the amplifier output as is the case with the feedback limiters described earlier—it would never open. It has to sense prior to the attenuator in the signal chain. This arrangement is called *feed-forward side-chain sensing* and is the prevalent method of generating control voltages in today's dynamics processors. Fig. 25-82 shows a typical gate circuit using this method; the input signal as well as going to the attenuator hits a variable gain amplifier, which determines the threshold. The more gain in the amplifier, the sooner the detector threshold is reached. Following the threshold detector—which is in this case a comparator type yes/no level sensor—are the various time constants. Depth is controlled by placing a limit on the amount of attenuation possible.

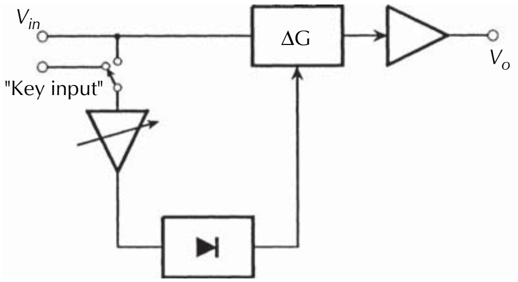


Figure 25-82. Gating feed forward sidechain processor.

25.12.2.2 Subtractive Feedback Gate or Expandor

An alternative gating method cunningly uses a limiter, albeit with a very low threshold, to subtract from or cancel the straight signal, Fig 25-83. The signal gains through both the straight path and the limiter (below its threshold) are arranged to be the same but out of phase; they cancel out. Above the limiter's threshold the limiter output remains fixed but the straight signal is left unhindered, so the two no longer cancel, leaving the straight signal predominating. Time constants of the effective gate are determined by those of the limiter, threshold by the gain of the amplifier within the limiter loop, and depth by contriving a mismatch between unlimited level and straight path level to produce less than total cancellation and some residual.

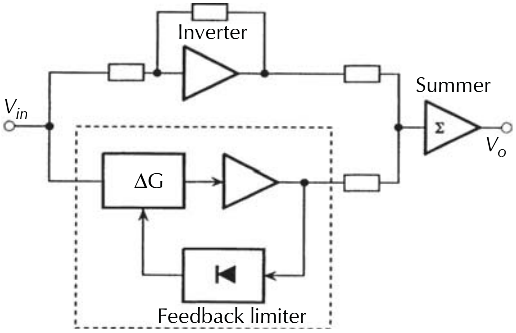


Figure 25-83. A subtractive gate, canceling a limiter from inverted input.

25.12.2.3 Keying

Keying is the triggering of a gate from an external source that is not from the signal that is actually passing through the gate. Perhaps the best and most commonly heard examples are keyed snare and kick drum sounds—all the rage in the dark ages of disco. In circumstances where a new drum sound was needed or alternatively the existing drum sound was not fit for human consumption (very common on live stages), the existing drum sound is used to key a gate that is carrying in its signal path something which can be convincingly shaped into a better drum noise. Favorites are white, or EQ'ed white, noise to emulate a snare sound; similarly, some tone around 20–60 Hz (sometimes even ac line hum) when shaped by the attack, hold, and decay times of the gate can make for a good kick drum!

25.12.3 Compression

As briefly outlined at the beginning of this section, compression is where the output signal from the processor does not increase as much as the input signal is increasing. If an input signal jumps in level by 10 dB, a compressor with a ratio of 4:1 would only allow the output to rise 2.5 dB. Correspondingly, a drop in input level of 16 dB into the same compressor would result in 4 dB output level change. A compressor reduces the dynamic range of an input signal by the amount of its ratio.

A true compressor acts on all signals, regardless of actual signal level, in the same manner. No matter if the input signal is way down at −60 dBu or up at +20 dBu, a change in input signal level of a given amount will cause a similar, reduced, change in output signal. Practically speaking, there is no such thing as a true compressor; things that come close and work down to very low signal levels are used in noise reduction systems for telephone lines, tape recorders, and wireless microphones, where they are used with a complementary expander (see later) to reinstitute the original dynamic range.

Most compressors have a threshold below which they leave a signal unscathed (a 1:1 ratio) and above which they proceed to compress the dynamic range, much as a limiter does. The family resemblance becomes all the more striking as compressors with high ratios are considered. A 10:1 compressor above its threshold reduces a 10 dB input level jump to just 1 dB. Infinity-to-1 reduces anything above the threshold to the same output level. Looks like a limiter, smells like a limiter. Generally, compressors are used at far gentler ratios (between 1.5:1 and 4:1) to bring up lower level program material in a less take-it-or-leave-it manner than a limiter while leaving some sense—albeit reduced—of light and shade, louder and quieter. They are also used to subtly make sounds chunkier—a degree of compression tends to accentuate lower frequencies, which are those generally most predominant and are so controlling of the gain reduction.

Major differences between limiting and compression are in the nature of the side chains, the level detectors, and in particular typically applied time constants. Limiters almost invariably have peak detectors, such that the peak of a waveform is detected and, time constants allowing, protected from overload by the limiter; this is fine for the protection mandate of limiters. Compressors, on the other hand, tend to have much more relaxed attack and release times, such that they are less intense sounding than the typically frenetic limiter settings. Similarly, since the concern with compressors is less peak level than loudness to the ear, which tends to gauge by overall signal energy or power rather than peak values, the detectors are typically average or power sensing. The slower time constants go a long way toward this by essentially ignoring peaks and responding more to an average over the time imposed by the attack and release times. Deliberate averaging detection (as opposed to the more or less accidental) sounds much more even and unobtrusive than peak detection; taking it a step further, power detection by means of a root mean square (rms) level detector is better yet. In reality, though, there's little to choose between average and rms detection, since although they give significantly different answers under test waveform circumstances, dancing around in the heat of audio battle they are quite difficult to tell apart. Occasionally either one or the other will be fooled by a difficult piece of program material.

The threshold tends to extend farther down for a range typically of −30 to +10 dBu, and the ratio adjusts from 1:1 (straight) usually to infinity:1 or close (limit). Some commercial units extend the ratio beyond infinity to negative values; that is, if a signal progressively exceeds the threshold it gets progressively further attenuated! Although on first glance it seems a bit pointless, it does allow fairly nice sounding level control for a compound signal; it permits looser (longer) attack times than would be possible on an ordinary limiter, with the resultant overshoot merely propelling the signal farther downward away from possible overload. It is also good for some pretty silly effects on individual instruments.

25.12.4 Expansion

An expander increases the dynamic range of its output signal in relation to the input signal. Its ratio determines how much: a 1:3 expander renders a 4 dB level shift in input signal, which results in a 12 dB difference at the output.

As mentioned under compression, true full-range expansion is a rarity and is generally only found as a complement to a compressor in a double-ended noise reduction system. In these circumstances they are nearly always of 1:2 ratio with an axis point (the level at which the input signal is the same as the output signal level) of around 0 dBu.

Practical expanders come with a threshold setting, above which they leave the signal alone and below which gain reduction sets in. Sounds a bit like a gate? A gate can be emulated by an expansion with a ratio of 1:infinity; any signal below the threshold gets attenuated away completely—the one exception is that expanders usually don't have a depth setting. The purposes of expansion are very similar to those of a gate, only generally it can sometimes do a better, less noticeable, job. A relatively gentle expanding slope (say 1:2 or 1:3) can provide the same degree of noise reduction as a gate with less abrupt changes in gain; since the signal is audible still (but quieter) and doesn't have to be resurrected with a start to normal level, fairly gentle (slower than a gate) attack times do not have as noticeable a softening effect on the required leading edge.

Expansion side-chain time constants are similar to those for a gate, as is the threshold range. Ratio, as with compressors, is usually 1:1 to 1:-infinity, although often "classic" implementations have a fixed ratio of something close to 1:2. Expansion is used as subtle gating in much the same way as compression is a gentler substitution for hard limiting.

25.12.5 Feed-Forward VCA-Style Dynamics

The feed-forward class of dynamics owes itself to the development of VCAs and similar log/antilog processing; it is exemplified by the classic dBx160 series. As far as consoles go, the mere existence of a VCA in the channel for fader automation begs for this style of processing to be incorporated. (VCAs are further discussed under Consoles and Computers, later.) Figure 25-84 shows such a processor in block diagrammatic form.

Key to VCA dynamics is the inherent exponential (logarithmic) control, which relies on reasonably simply implemented basic transistor behavior (base voltage versus current). Gain (or gain reduction) of a VCA is as good as linear dB/v, which can lend to a deterministic design approach (meaning one can pretty well predict what the circuit will do within narrow limits, without a servo loop to help). Simple log/antiloging lends itself to another typical feature of VCA dynamics sections.

25.12.5.1 rms Detection

Hitherto, detection of signal levels in dynamics had been either peak or average. These were actually achieved by broadly similar circuitry with the difference dictated by the attack time applied after signal rectification; short attack times allowed the reservoir capacitor to charge immediately to the highest signal level applied, while longer attack times tended to smooth out the peaks, settling on an average value of the applied rectified waveform. And a sort of mushy continuum existed between the two.

Rms (root mean square) detection has the intent of providing a measure of the energy in an applied waveform, the actual power. The reasoning is that a power measurement could be considered more equivalent to loudness. Rms is achieved by first squaring the applied signal (i.e., multiplying it by itself, not turning it into a square wave), finding an average value of that squared source, and then determining the square root of that average (unsquaring it). Seems like a lot of bother to go to, doesn't it? Well, no one would have bothered if there was a reasonably straightforward method. This comes from an application of a log anticell.

A precision-rectified (meaning accurate down to very low levels) input signal is logged, and then its output is doubled (added to itself); doubling a log value squares the number it represents. This signal is then integrated with a time constant long enough to allow reasonable averaging of the lowest frequency under consideration; this incidentally defines the minimum attack time of the processor. This log value average is then halved (division of a log value by two is the same as finding the square root), delivers a log-world rms-detected output. (In this circumstance a subsequent antilog conversion is unnecessary. Actually, the square rooting is ignored at this point, too, since it can be achieved in a later scaling exercise.) The good news is that all that can be done with a handful of transistor junctions. Release time can be extended with a following buffered capacitor, but often the imbued time constant of the rms detection serves as symmetrical attack and release. This somewhat leisurely time response (necessary to permit good rms detection at low frequencies, devoid of distortion-creating ripple) in and of itself ensures that the behavior of such a dynamics section can't get too wild and interesting, but by corollary such processors do afford probably the least intrusive method of automatic volume control, which is a highly prized attribute on occasion.

25.12.5.2 Thresholding

The rms-detected control signal is then masked in a threshold determining circuit; typically this is a precision rectifier with its reference point determined by a threshold control voltage—the purpose of this is to ignore all variation of the detected voltage until it exceeds (in the case of a compressor) the threshold point, beyond which its output follows the rms detector output. Any control signal escaping the thresholder still has a dB/V characteristic, being still logged, following the input signal. Another (linear-think) way of looking at this is that a division takes place; the detected control signal is divided by the threshold, but with any result less than 1 masked out at 1, only greater than unity results being passed.

If the thresholder is designed to pass only changes below the threshold, then the low-level effects of expansion and gating are possible, signals above threshold being ignored. Separate thresholders and following conditioners are necessary for each desired function of the dynamics section.

25.12.5.3 Ratio

If this thresholded control voltage were applied directly to the (level-adjusted) control port on the VCA, something odd would happen: nothing. More precisely, above the threshold the control signal would rise in accord with a rising applied audio signal to the precise extent that the gain reduction resulting from it would be exactly the same as the increase in signal. The VCA output would remain at a fixed level for any applied signal level above the threshold. In other words, it is a compressor with an infinity:1 ratio, meaning that above the threshold, any amount of signal level variation will have no effect on the output. In yet other words, it would be a limiter (albeit with slow dynamic response).

Introducing a variable attenuator in the feed to the VCA control port from the thresholder affords altering the amount of dynamic gain reduction; less control signal variation, less gain reduction. A nice feature of this very simple approach in log world is that this attenuation (equivalent to solving for variable roots in linear) results in precise applied signal level to dB gain reduction ratios; for a given setting, if the input signal were to rise 6 dB, the output would rise only 3 dB; this ratio, 2:1, would obtain linearly for any applied signals above the threshold.

Astute readers may wonder what would happen if the control signal, rather than being attenuated, was

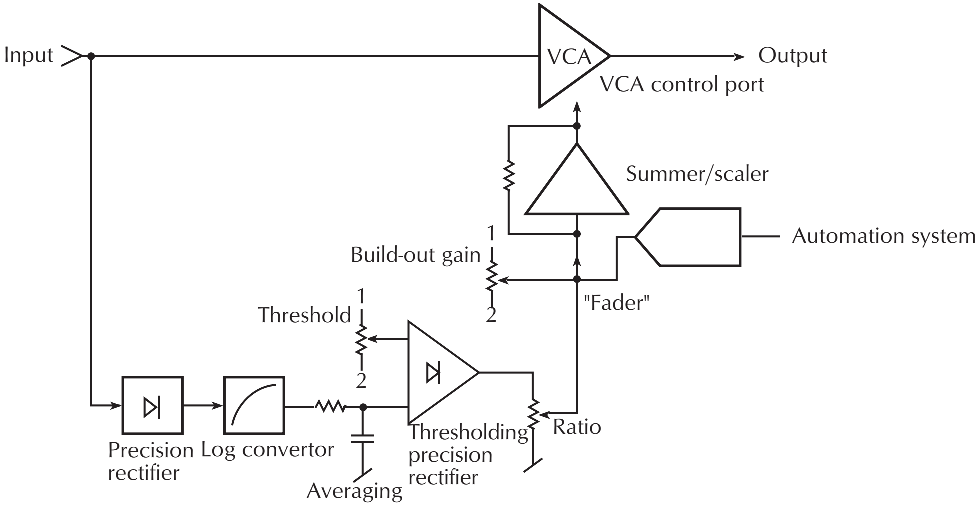


Figure 25-84. Block diagram—feed-forward VCA Dynamics.

amplified instead. Yes, the degree of attenuation would become bigger than the corresponding changes in applied signal, and the VCA's output would actually increasingly *reduce* beyond the threshold. This effect was brought to stardom by the Eventide Omnipressor.

25.12.5.4 Side-Chain Sources

Figs. 25-84 and 25-85, the block diagram and schematic respectively, of a feed-forward style dynamics element, both only consider the side chain as being taken from the same place as the input to the VCA gain controller. This need not be the case at all. In fact, with a standard channel's signal processing it would be quite a limitation, forcing the dynamics section to be solely post everything, prefader. The side chain input may be separated and instead taken from pretty much anywhere upstream; the overall effect is (with only some odd side effects) the same as physically moving the whole dynamics section to that location. For the most part, the audio doesn't care that the control voltage from the sidechain has no relation to that being passed through the VCA. The main areas where things might sound awry are if there is a significant (and audible) deliberate time delay between sense and activation or extremely short time constants are invoked. The first disconnect is obvious, and the second is almost irrelevant since it wouldn't be sounding very nice anyway.

Taking this a step forward again, a prime virtue of a VCA-based system is that many different sources can operate simultaneously on the one relatively expensive VCA gain element. If there are multiple side chains (say, one for keying/gating/expansion, another for compression) these again need not sense from the same pick-off points but places more suited. The gate would likely sense postinput filters, while the compressor would likely be farther downstream, one side or the other of the EQ. This situation would work, but would result in behavior unlike having a discrete gate up front and a discrete compressor downstream. In the literal case—i.e., with two separate dynamics elements—the incoming signal would be gated before it hit the compressor. In this virtualized case, though, the actual audio signal hitting the compressor side chain would not have previously been gated and hence cause the compressor to act differently than if it had. Subtle, maybe, but a definite difference.

25.12.5.5 Makeup Gain

The side chain's thresholded and ratioed control signal is summed in with a voltage representing the amount of buildout gain (necessary to compensate for the signal level reduced by the effect of compression/limiting) and is also, in the case of a typical console channel, summed in with the voltage from the automation system representing the fader position. This summation is scaled to suit the actual (highly sensitive) VCA control port, to which it is fed.

25.12.6 A Practical Feed-Forward Design

A highly integrated part, the THAT Corporation's 4301, has both an rms detector and a VCA built in, in addition

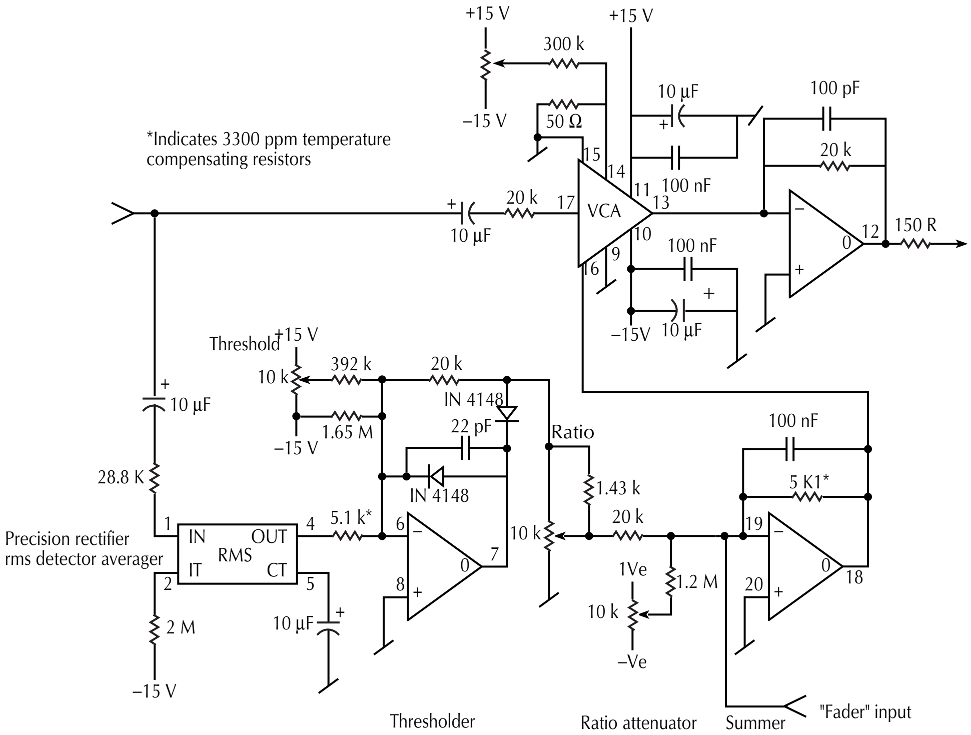

+15 V
+15 V
300 k
−15 V
50 Ω
10 µF
100 pF
*Indicates 3300 ppm temperature compensating resistors
20 k
15
14
11
100 nF
13
17
VCA
10
9
16
100 nF
12
150 R
0
10 µF
−15V
10 µF
Threshold
+15 V
10 k
392 k
20 k
IN 4148
22 pF
1.65 M
−15 V
Ratio
100 nF
5 K1*
+
10 µF
28.8 K
IN 4148
1.43 k
20 k
1
IN
OUT
RMS
IT
CT
4
5.1 k*
6
7
10 k
19
18
20
Precision rectifier rms detector averager
2
5
8
1Ve
1.2 M
10 k
−Ve
2 M
10 µF
−15 V
Thresholder
Ratio attenuator
Summer
"Fader" input

Figure 25-85. A feed-forward VCA compressor using THAT 4301.

to some op-amps for glue. A simple, very low parts count compressor with all the active elements contained within this part is shown in Fig. 25-85. As can be seen, it relates strongly to the block diagram of Fig. 25-84.

25.13 Mixing

25.13.1 Virtual-Earth Mixers

The circuit diagram of Fig. 25-86 in its simplicity belies the hidden design that is in the relationship of the circuitry to its mechanical and electrical environment.

This is where the care and feeding of op-amps (Section 25.7) and grounding paths (Section 25.8) really pay dividends. Mix-amp stages, with large numbers of permanently assigned sources such as in the main mix buses, are as crucial to the overall well-being of a console as any front-end stage could be. In a typical situation, as a unity-gain virtual-earth mixing stage with 33 sources (channels plus access), the amplifier is being asked for about 30 dB of broadband gain, as much as any other stage in the chain including both the microphone preamp and/or secondary input stage.

25.13.2 Noise Sources

All the following about mix devices assumes that system grounding is impeccable. Jolly good. That said: That mix-amp gain is sometimes referred to as *noise gain* is not accidental. Unless care is taken to balance fader-back channel noise contributions against this self-generated mix-amp noise, the latter could well predominate and arbitrarily determine the noise floor for the entire console. Similarly, channel noise contribution should equal or outstrip mix-amp noise, but not excessively so: ideally they should equally contribute, to the extent that channel-off noise contribution should not necessarily impact the overall bus noise, while bus noise should not significantly impact channel-on noise. Self-noise generation in the mix-amp is predominantly the amplified thermal noise of the paralleled source and feedback resistances, device input current noise, and surface generation and recombination noise. The last two can be minimized by device choice. Thermal noise is physics and is here to stay. Common sense on first glance says to make the mix resistors as low in value as possible but this has the downside that too low a value would cause quite large signal (hence, ground) currents

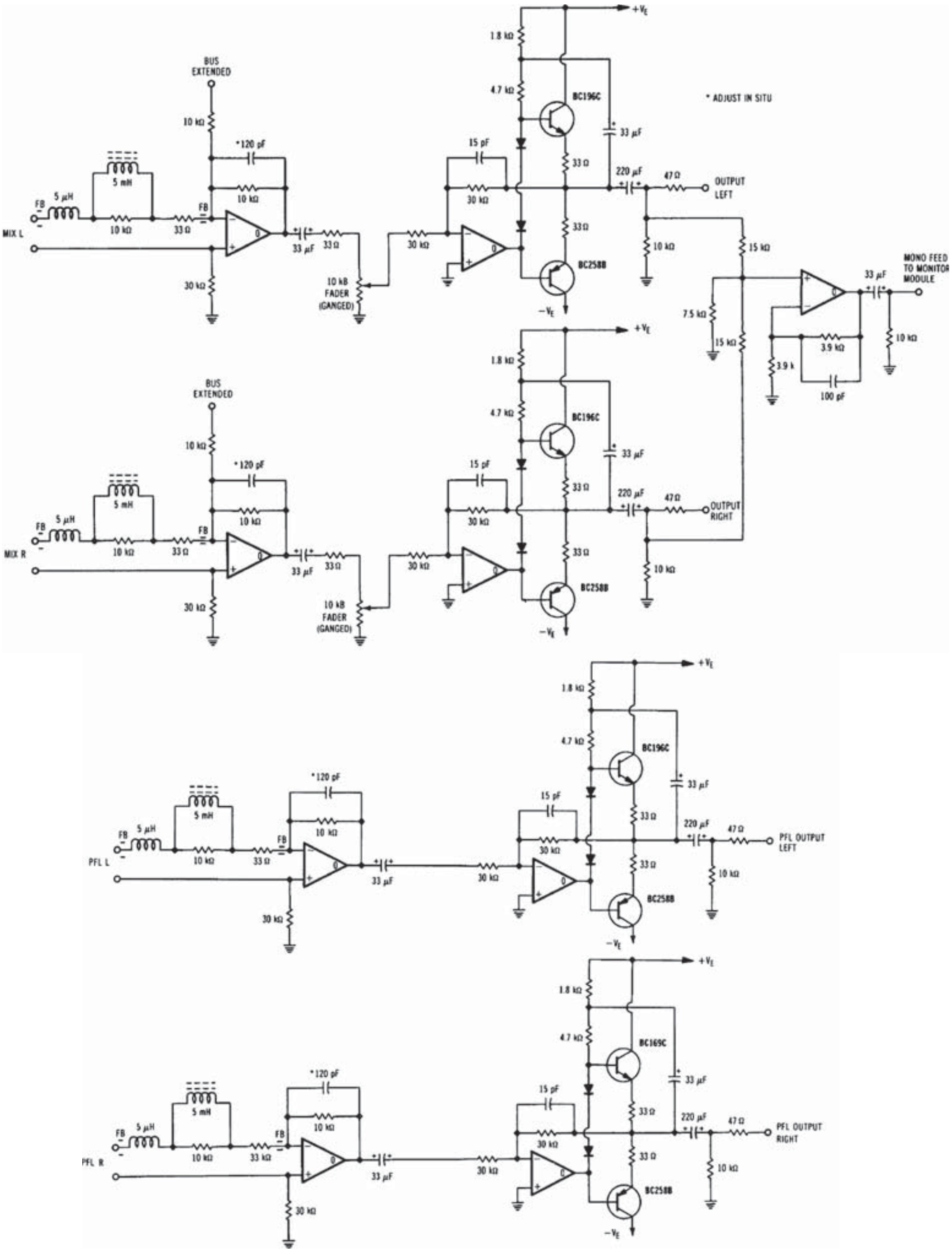

+V_E
1.8 kΩ
4.7 kΩ
BC196C
* ADJUST IN SITU
BUS EXTENDED
10 kΩ
*120 pF
15 pF
33 µF
33 Ω
220 µF
47 Ω
OUTPUT LEFT
5 mH
10 kΩ
30 kΩ
FB
5 µH
MIX L
33 µF
10 kB FADER (GANGED)
BC258B
−V_E
15 kΩ
MONO FEED TO MONITOR MODULE
7.5 kΩ
3.9 kΩ
3.9 k
100 pF
OUTPUT RIGHT
MIX R
PFL L
PFL OUTPUT LEFT
BC169C
33 kΩ
PFL R
PFL OUTPUT RIGHT

Figure 25-86. Master group electronics.

to be thundering about. On a less technical and more economic level, it necessitates somewhat beefier and more serious buffer amplifiers on each source to feed the buses.

Ordinarily though, the mix resistors are of such a value that, in the context of a complete mixer, the combined effectively paralleled resistance is well below the optimum source impedance of nearly any mix-amp device available, so the primary noise modes are those above-mentioned device vices. This certainly isn't too difficult with FET front-end devices, with their high OSI (optimum source impedance). These devices have a couple of other major benefits in this application though by virtue of their FET inputs. Input current (hence, input current noise) is extremely low, and being FETs they don't have the many low-frequency junction and surface noises inherent to bipolar devices. It seems a paradoxic absurdity to use an ultrahigh input impedance device for zero impedance mixing, but in many ways and under some circumstances they're better suited than bipolars. On the other hand, the intrinsically superior noise performance of a 5534-class device can pay dividends in this application. Like so many cases in console design each individual application needs staring at for its own optimum solution. This is all really only a problem for those who have the luxury of designing small mixers or where it is more or less guaranteed that only a small number of sources will be allowed to hit the bus simultaneously and hence where the parallel impedance of the sources remains fairly high. In most midsize and large consoles without these constraints, mix device noise will likely predominate. Device choice will be down to its self-noise (of course), and output current capability if the summing resistor value is low, and ability to cope with a big hairy capacitative bus sitting on its input current node. Integrated mic-amps have been successfully used as differential passive mix-bus amplifiers, which with their very low OSIs stand a chance of getting closer to that low bus impedance and low bus noise nirvana. However, as alluded to earlier, the channel-off noise contribution from all those bus-driving amplifiers in all those channels is more likely to then predominate. It is a balancing act.

If bus noise performance truly is a major concern (as it could possibly be on a tracking console) removal—as in physical disconnection—of all unused sources from the bus at all times is the best way to get the noise gain down and that bus impedance back up to where mix-amp noise can be optimized to it. No way to run a railroad or a mixdown console, though.

Things can get a bit startling if the resistance/OSI relationship is awry. Above the OSI as much as below its OSI, device noise becomes an increasingly important noise contribution. Many years ago in a mixer design with bipolar device mix-amps and quite high mix resistors, the measured bus noise was actually quieter on a 20-channel version than on the 10-channel original. It wasn't until much later that what was actually happening finally dawned. Increasing the number of source resistors reduced the bus impedance, previously well above the OSI of the amplifier with only 10 sources, to closer to the OSI, where input noise voltage was contributing less.

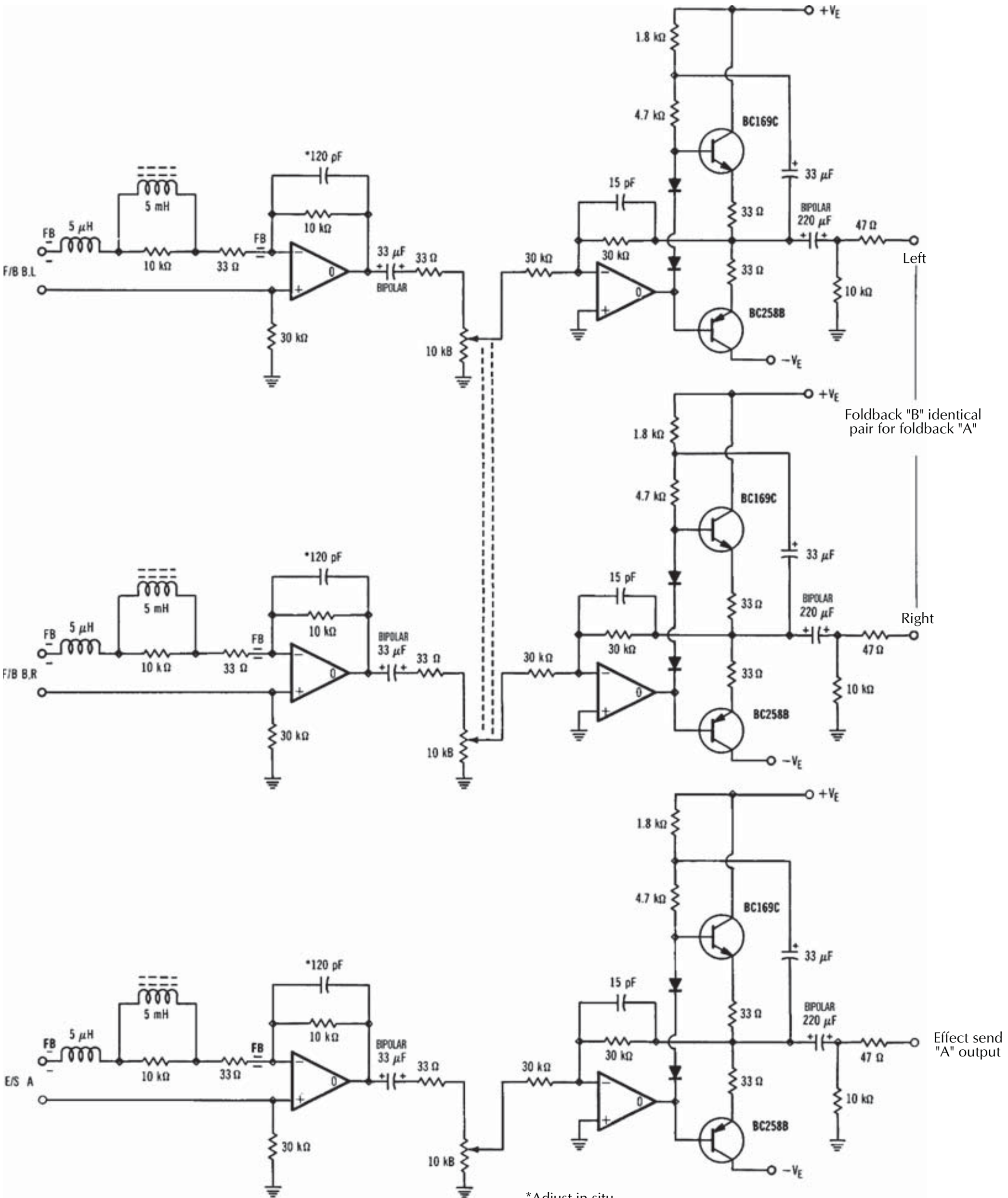
Theoretical source impedance and device contribution tell less than half the story in a practical design. They may be quantifiable in the isolation of a test bench, but thrown into a system they can all seem a bit meaningless. It's all largely a matter of grounding and out-of-band considerations.

25.13.3 Radio-Frequency Inductors

Inductors are used between the bus and the amplifier input in Figs. 25-86 and 25-87. A simplistic view is that they are there to stop any radio frequency on the mix bus from finding its way into the electronics, but this is only part of their purpose. The ferrite beads and small chokes (about 5 μH) are there to increase the input impedance and hopefully help decouple the bus from the amplifier at very high frequencies. The larger inductance creates a rising reactance to counteract the falling reactance of the bus capacitance. If left completely unchecked, this capacitance would cause the mix-amp extreme high-frequency loop gain to turn it into an RF oscillator. Feedback phase leading around the amplifier stops the gain from rising, but if it were not for some series loss (accidental or deliberate) in the input leg, it would be insufficient to hold the phase margin of the amplifiers within their limits of stability, especially at bandwidth extremes where device propagation delay becomes significant in the loop. A small series resistance can provide this loss while also defining the maximum gain to which the circuit can rise. A parallel inductor-resistor combination improves on this in a few important respects.

The inductor is calculated to present low in-band (<20 kHz) reactance, allowing the mix-amp to operate on the bus in its intended virtual-earth (zero-impedance) configuration. The reactance rises gently at the audio high-frequency end, imparting little frequency response anomaly but a definitely beneficial partial phase straightening against the inevitable effect of heavy bus capacitance.

At even higher frequencies, the inductive reactance continues to rise until the combined network impedance is limited by the resistor, which is of high enough value to define amplifier out-of-band gain to a reasonably low value. It is low enough, however, to stop the inevitable inductor-bus capacitance resonance from getting completely out of hand. Making a stable inductance-capacitance oscillator is one way of preventing


+V_E
1.8 kΩ
4.7 kΩ
BC169C
33 μF
15 pF
33 Ω
BIPOLAR
220 μF
47 Ω
Left
*120 pF
5 mH
10 kΩ
FB
5 μH
F/B B.L
33 Ω
33 μF
BIPOLAR
30 kΩ
0
10 kΩ
BC258B
−V_E
10 kB
Foldback "B" identical
pair for foldback "A"
Right
F/B B.R
E/S A
Effect send
"A" output

+3 Identical for effect sends, C & D

*Adjust in situ

Figure 25-87. Auxiliary group electronics.

spurious instability, but is not exactly the desired end here.

While FET inputs are far less prone than bipolar inputs to the intermodulation and direct demodulation effects that cause RF interference to appear out of nowhere, this fairly healthy brace of filtering may be helpful to those living near a source of high-powered very high frequencies, such as a group of television transmitters.

25.13.4 Virtues of Grounding

Grounding paths for virtual-earth mixing, especially in long mixers, are always the final arbiter on how far down the system noise floor will go and how susceptible the mix stage is to extraneous fields and earth currents. In this age of digits, ground paths are especially crucial. Remember from Fig. 25-65 how the ground noise on the noninverting input of an op-amp mix stage gets amplified up by the noise gain of the stage? This implies that a ground noise of –100 dBu will end up at about –70 dBu for a 32-source mixer, which is hardly adequate.

A simple, but so often ignored, rule with virtual-earth stages is to make sure that the ground reference has got the same dirt on it as the signal and vice versa. Yes, ground follows signal. If both ground and signal have the same noise in the same phase, there is a chance that the noise will get ignored as common mode and not amplified in the mix-amp. So, for each mix bus, there should be a parallel ground bus being fed by the last relevant ground reference from each channel. Avoiding a major bus-length ground loop (otherwise known as a single-turn transformer!) means that all the heavyweight signal current in the channel proper (e.g., fader/mute/mode switchers) has a direct wire to central ground while the mix-amp has a respectable output referenced ground to work against, clean of channel signal currents but representative of the reference of the buffer amplifiers. The mix-amp does not take a direct system central ground of its own.

25.13.5 Passive Mixing

There are, of course, alternatives to single-bus virtual-earth mixing. Passive resistor mixing, Fig. 25-88, is quite viable for fixed-assignation systems that are not going to be chopped, changed, or switched in and out. A major advantage is that bus capacitance is merely something to be taken into account in terms of frequency response and phase rather than directly imperiling the stability of the mix-amp. For passive mixing, the mix-amp is just a buffer amplifier to make up the loss in the resistor tree; RF filtering becomes simple with known filter source and load impedances together with the ability to refer against ground. A primary weakness is that the bus is unbalanced and is of some impedance at audio (albeit fairly low due to paralleled sources). As such it lays itself wide open to induced noise and capacitatively coupled crosstalk. Despite this, it is a method used with considerable success for many years in quite a few production mixers.

In all cases but especially in small mixers, say with fewer than eight sources, there is a theoretical noise advantage to passive mixing over virtual earth. As an extreme example, simple summing of two sources of passive mixing calls for 6 dB of gain to make up for the loss in the summing network. A virtual-earth mixer needs around 10 dB. Beyond eight sources this advantage tends to the insignificant.

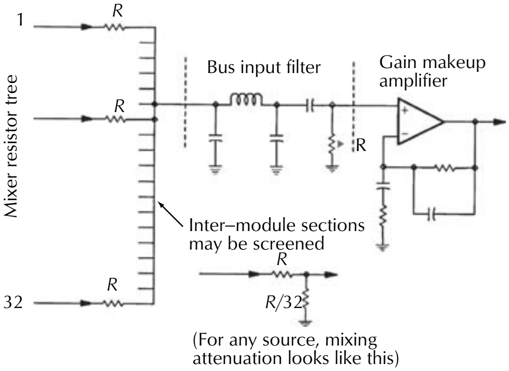

1
R
R
R
32
Mixer resistor tree
Bus input filter
Gain makeup amplifier
R
Inter–module sections may be screened
R
R/32
(For any source, mixing attenuation looks like this)

Figure 25-88. Passive mixing arrangement.

25.13.6 Devolved Mixing

Distributed or devolved mixing, Fig. 25-89, uses local mix-amps to sum relatively small blocks of channels; the outputs of these local amplifiers is then taken to a common summing point. This quite neatly obviates having to deal with a long bus but does create a practical problem of locating the distributed summers.

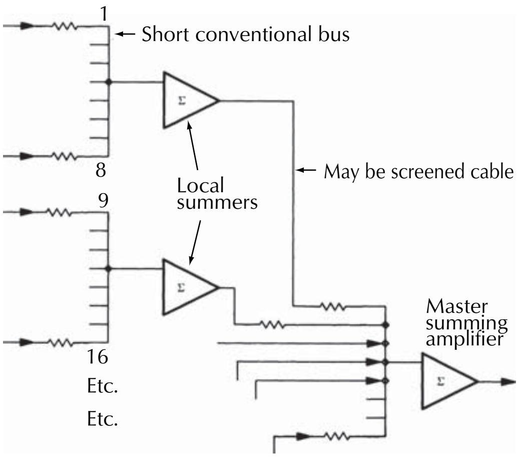


Figure 25-89. Distributed or devolved mixing.

Both passive and devolved systems have the advantage that large amounts of the bus can be run in shielded cable. The extra capacitance here does not have the awful consequences it does with long virtual-earth summing amplifiers.

For consistency—if this approach is taken—all buses should be run devolved. This means the submix facilities for the PFL buses, effect sends, foldbacks, the main stereo/monitor mixer, and analog subgroups (if used). Also, provisions must be made to arrange the master mixer for each of those at the grouping end.

25.13.7 Balanced Mixing

The earliest form of signal mixing consisted of directly paralleling the sources, which were generally medium-impedance (nominal 600 Ω) and balanced. This form of passive balanced mixing persisted until semiconductor electronics and its readily achieved zero impedance transpired. The balancing was done entirely by transformers; again, things that have fallen at least partially by the wayside. As a technique it was simple (for the technology at the time) and maintained all the advantages balanced systems have in general—principally a welcome robustness and immunity to interferences, induced noise, or crosstalk.

Balanced or differential mixing has became practical again with falling component costs and the development of simple electronic differential and floating balanced input and output circuits (see Sections 25.9.6.1 to 25.9.6.4). Fig. 25-90 shows how differential sources of the trivial kind (straight and inverted) can be mixed onto a balanced virtual-earth mixing bus, created and sensed by a superbal input stage.

Although requiring a comparatively large number of parts, the performance of such an arrangement in the context of a large multitrack console is truly staggering, especially noise, head room, electromagnetic field rejection, and crosstalk. The noise improves in two respects:

1. No longer is the mix-amp amplifying the noise on its reference ground. It is referenced to itself, effectively.
2. Square law noise summation—twice the signal (coherent) means 6 dB gain, two lots of incoherent noise 3 dB gain, bingo, 3 dB noise advantage.

Head room, by virtue of two signal paths carrying the same information differentially, is 6 dB higher. (Naturally the noise and head room are interrelated; whichever is more pressing in a given circumstance necessarily takes precedence in the level architecture.) The RF field and crosstalk rejection improvements are dramatic, but they really ought to be expected from the naturally self-canceling nature of balanced systems.

All the problems of keeping virtual-earth mixers tidy and stable apply twofold here; of course, bus buffering is strongly recommended, mostly to allow the bandwidth definition around the superbal to be effective.

Passive balanced mix-amps can be arranged around integrated mic-amp devices such as the THAT 1510; being single-ended output doesn't lend them to dynamically generating differential virtual–zero impedance mix buses, but does allow the choice of mix resistor values versus mix width to optimize the parts mix-noise contribution. It is presently difficult to consider any serious large console design that doesn't use balanced mix buses.

25.13.8 Pan Pots

As outlined earlier pan pots are a means of positioning a monophonic image somewhere within a stereophonic image plane. About the simplest pan pot is shown in Fig. 25-91A where a pair of linear potentiometer tracks are complementarily wired; one goes up, the other goes down. All well and good and even the sums work out nicely; if the L and R outputs are subsequently remonoed the summed signal remains at the same amplitude regardless of pan pot position—center, either end, or anywhere in between. Subjectively, though, the image seems too loud at the peripheries (i.e., extreme left and right) and subdued in the middle.

Replacing the linear pots with a ganged log/anti-log pot (the log section wired upside down) performs much the same function but with a different law, Fig. 25-91B. If a signal is panned steadily right, the left hand output is steadily attenuated, leaving the right output fairly steady in level (in practice it shifts about 1 dB). The center position sees both L and R only attenuated slightly (<1 dB) with respect to the starting mono signal. Not surprisingly this has the opposite subjective effect to linear pots: the image seems louder in the middle than at the peripheries. Despite that, often this law is more appropriate, particularly where the pan pot is used as part of multitrack odd/even panning or in use as a correctional offset control. In these cases there is virtue in leaving at least one side fairly unscathed.

Somewhere between these two extremes (if extremes is the right expression for a 6 dB difference) should lie a happy medium at which the signal keeps an even subjective level panning across the image plane and also tracks well (i.e., has good correlation between control position and image position). Easy? This has been the

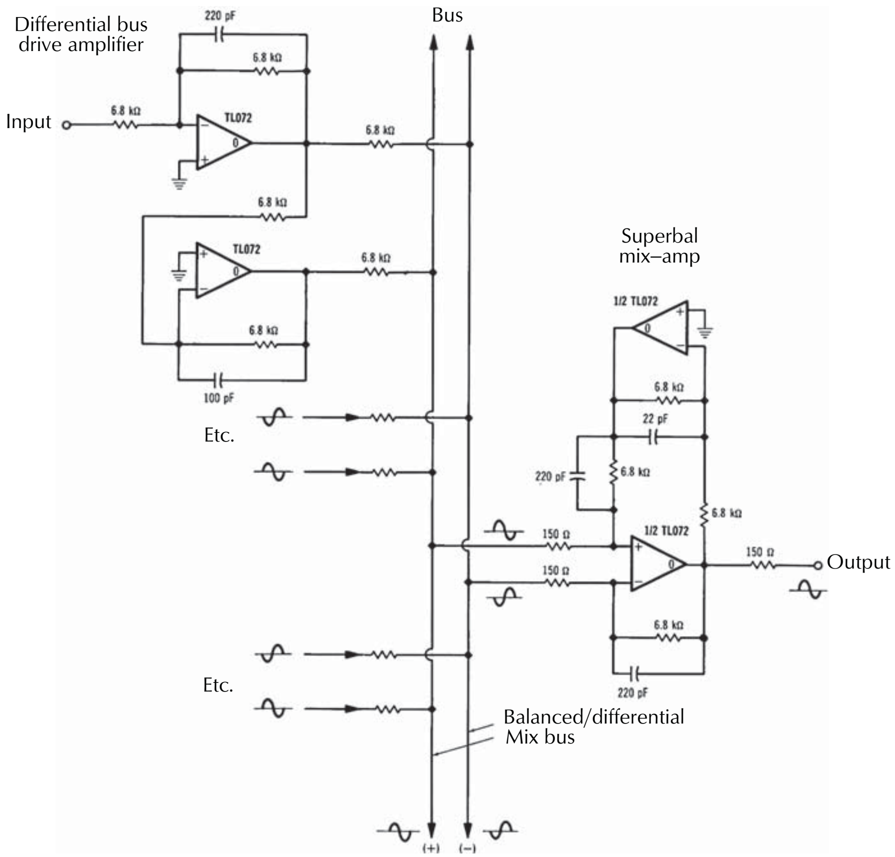

Differential bus drive amplifier
220 pF
6.8 kΩ
Input
6.8 kΩ
TL072
6.8 kΩ
Bus
6.8 kΩ
TL072
6.8 kΩ
6.8 kΩ
100 pF
Etc.
Superbal mix–amp
1/2 TL072
6.8 kΩ
22 pF
220 pF
6.8 kΩ
6.8 kΩ
150 Ω
1/2 TL072
150 Ω
150 Ω
Output
6.8 kΩ
220 pF
Etc.
Balanced/differential Mix bus
(+)
(−)

Figure 25-90. Differential balanced mixing.

subject of raging controversy and opinion for decades. Should it be 2, 3, 3½, 4, or 4½ dB down at center?

As is often the case, those closely involved with the theorizing somewhat lost touch with how a pan pot is ordinarily used. A pan control usually remains rusted at an initially set position for hours, days, weeks, months—however long the mix takes. If a pan pot is used dynamically for effect during a mix, its very drama drowns any question of whether it was "a wee bit quiet in the middle."

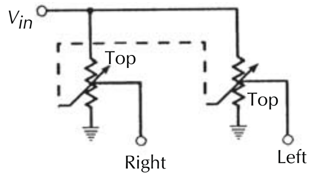
A single pot used as in Fig. 25-91C allows a choice of central down points by adjusting the relative values of the source resistors and pot value, but at the cost of slightly iffy tracking (most pan effect tends to happen at the extremes of the control travel) and ultimate panning. When panned hard one way it is nearly impossible—due to wiper-track resistance—for the diminished side to achieve complete attenuation. If 40-odd dB is good enough then this may be the one. Shown are values for a 3 dB down panpot.

During the 70s the BBC had the understandable problems of multitudinous operators, countless consoles of varying antiquity, and a considerable number of console suppliers. Their aim was consistency and to this end evolved a dazzlingly simple arrangement shown in Fig. 25-91D recommended for inclusion in new supplied equipment; it works.

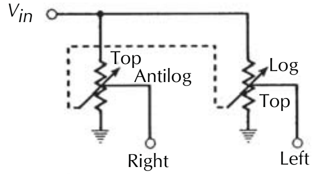
25.13.9 Surround Panning

Surround has many variants, but for the purposes of discussion here 5.1 will be considered; other formats have basically similar requirements that may be taken as a subset or extrapolated from 5.1.

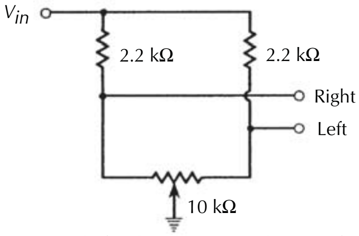
Well, quad is back (and it hasn't forgiven). At least in terms of four of the 5.1 signal paths, left front, right front, left rear and right rear. Panning for these is achieved in much the same manner as it was in those chillingly far-off days, a joysick that controls relative proportions of the source signal to the four output, or a



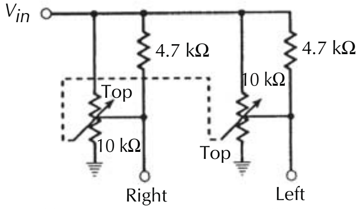
A. Ganged pair of counter–connected linear pots.



B. Pair of ganged log/antilog tracks as routing panpot.



C. Single pot with feeder resistors. Choice of resistors gives choice of center down point, as shown approximately 3 dB.



D. Ganged linear track pair with pull–up resistors approximately 3 dB center attenuation.

Figure 25-91. Panpots.

pair of pots that do substantially the same; one pans left/right, the other front/rear. The "1" of the remaining 1.1 is center; it is to where central dialogue or vocal is panned; this is usually achieved with a blend or similarly named control, which cross-fades the source signal between the center channel and the quad pan pot. In this manner, a source can be directed to any one path, all, or a combination for desired effect.

The last 0.1 is a bass sublow channel. The .1 means that it is (but not always) band limited. It usually has its own level control independent of the full-bandwidth panning.

The surround panned outputs form a channel with six dedicated surround mix buses, which are treated as a married set within the console much as the main stereo mix-bus is/was.

25.14 Monitoring

Monitoring is probably the single most important section of a console. Without it the engineer cannot listen to the results of his labor. At its simplest, monitoring consists of a power amplifier and loudspeakers hung across the main output(s) of the console, with the auxiliary functions either unused or preset. In public address (PA) work the PA actually is the monitoring; the only other function necessary is prefade listen (PFL) and then really only during panic mode. At an alternate extreme the monitoring demands for multitrack recording extend to an entire secondary submixer replete with panning, pre/post foldback effect feeds, and stand-alone soloing, together with listen access to all console send and return ports. The in-line console principle makes efficient use of electronics to combine often coincident signal and monitoring path requirements for normal multitracking techniques. If the architecture is well thought out, it is operationally rare to need to listen to anything other than the main stereo bus output; this output serves as both the multitrack monitoring bus and the stereo mixdown bus.

Three distinct types of monitoring activities evolve in multitrack work:

1. Mainline—The stereo bus encompasses the multitrack machine sources/returns and stereo mixdown. This can be read as surround bus if appropriate.
2. Transient—This allows short-term check listening of individual channels for reassurance or adjustment, using PFL or solo functions.
3. Auxiliary—This provides access to the assorted foldback/effect feeds, effect returns, mastering machine, and subsidiary machine returns.

From an operating point of view, the foregoing activities seem to form natural divisions. From a technical stance, it's a different matter entirely. The solo (*in-place monitoring*) function is very closely related to the stereo bus. In fact, it uses exactly the same signal path throughout—and can be seen simply as a modified use of it. PFL, though, despite a similar operation (only prefade as opposed to post pan listening), actually requires an entirely separate bus and mixing system. Its output is switched to override the main path into the monitors. (It may seem strange to go through all this for a spot-check function that tells less than the stereo in-place solo, until it is remembered that a solo disrupts the mix while a PFL is nondestructive.) Conversely, an operator usually has a psychological hook about the main stereo bus monitoring being the gospel unblemished signal path and that all the auxiliary functions are somehow less polished and somehow tainted. In reality, the monitoring chain normally selects directly between all its sources, merely treating the stereo mix as one of the many. No special treatment is desired or given.

25.14.1 Solo, Solo-Free, and Prefade Listen

An assumption is made that the solo function is such that if a console channel is soloed, all other sources contributing to the main stereo bus are muted, leaving the desired channel in isolation at its set level and panned position. An exception and extension to this are for other channels (principally those returning effects to which our soloed channel may be contributing) to remain unmuted in the stereo mix during solo operation; this is done by using the solo-free button on those channels still needed. Solo-free detaches the channel from the consolewide muting/solo activation logic.

Soloing individual channels wet (i.e., with all its attendant effects) is a common need; at a stage in a production where things are dripping in reverb and sundry funny noises, soloing in context only makes sense—by that time it is well known and redundant what something sounds like dry. A channel's sound has become an amalgam of the source and applied effects, not just that of the source.

The upshot of this is that solo monitoring is inherent to the stereo mix path. If that path isn't selected for monitoring, then neither is the solo. So, although a solo overrides the main stereo mix (unless disabled altogether by a master function, solo safe), it cannot override anything else, unlike the PFL.

Although PFL could just be brought up as another monitored source, it is made to emulate solo in single-button touch operation, with the added advantageous capability of overriding everything—whatever is selected to monitoring. Hit a PFL button anywhere on the console and, if desired, it will be what you hear in the monitors. Alternatively it can be arranged to just come up on headphones or a "near field loudspeaker" so as not to disturb the main monitors.

25.14.2 Monitoring Controls

Now we've worked out how to get what signal and at what priority into the monitoring chain. What other torture do we put it through?

1. Level control, which is used to adjust the volume. Usually a big knob or a fader of its own. *The* most used control on any console—just ask any console manufacturer's service department.
2. Mute is used to turn the row off occasionally.
3. Dim is used so that you can hear what people say.
4. Mono is still used in radio and TV.
5. Phase reverse is used to make sure you haven't already done it inadvertently. (This function together with the mono button makes for one of the quickest ways in history of lining up analog tape machine azimuth.)
6. Split is unashamedly borrowed from broadcast monitoring technology. This routes a mono sum of the main stereo mix bus continually to the left side of the monitor chain and a mono sum of whatever source is selected (including PFL override) to the right side, providing simultaneous monitoring of two different sources—one of which would almost certainly be console output anyway. (Split's origins lie in network radio, where announcers on the air have to talk up to program junctions and smoothly hand over to another studio or network feed, news, or whatever at a cue. In order to do this, they have to be able to hear both themselves and the network they are opting into to hear the lead-up and handover cue.) Other than its primary design use, the split function is used considerably under other normal programming, affording random source monitoring without losing track of what the main console output is doing. It's also used extensively in program prerecording and production, enabling, with practice, real-time multisource edits (jump edits) without recourse to razor blades and tape. Split will eventually find a niche in multitrack recording techniques; if nothing else, it can fulfill the requirement for single-loudspeaker mono monitoring, by simply selecting the right side to a dead source.
7. Desktop loudspeakers, or idiot speakers, are used to do transistor-radio and cheap hi-fi impersonations, also affording a respite of sorts from the sometimes wearing grandiosity of normal monitor loudspeakers.
8. Near-field loudspeakers (relatively small speakers, usually perched on the console's meter bridge) are used as a twofold reality check during mix: they are close enough to the engineer for the room acoustics to be unimportant; they are closer in size/quality to what the majority of listeners will be using. Often, they are used as the prime monitoring with big

monitors and idiots being used briefly to make sure nothing's gone amiss or become overblown.

25.14.3 Related Crosstalk

In a program sense, two forms of crosstalk are relevant. The first, related crosstalk, is a signal bleeding over into another signal path that is carrying a musically and temporally related signal (e.g., between the left and right of a stereo pair or between adjacent tracks of a multitrack recorder). It happens quite often and is fortunately not often subjectively obvious or embarrassing; usually they're playing the same song!

Crosstalk within multitrack recording systems is usually little short of horrifying. As a result of the large physical size of the console, ground paths are unavoidably long and ground currents generate (and cross-inject into other paths) crosstalk voltages across the resultant ground impedances. Capacitance between interconnecting cabling, looms, modules, buses, indeed everything, results in a reasonably suspect electrical overall crosstalk performance. Naturally, the better the design and construction, the better a console tends to be in this respect. One typically gets what one pays for.

This was overshadowed and mitigated by analog multitrack tape machine crosstalk between tracks—a safe order of magnitude worse than even a horrid console could be. These tape machines not only had the same electrical problems as consoles but also had many magnetic heads in very close proximity, all dealing with a tape medium not notable for magnetic isolation anyway. It was all tolerable and usable simply because all the crosstalk was related and blended in unnoticeably.

25.14.4 Unrelated Crosstalk

Unrelated crosstalk is the clashing and cross-bleeding of signals that have nothing whatsoever to do with each other and are a mutual embarrassment.

In console monitoring a hostile signal (i.e. a delayed replay B check of a master) can be screaming about in uncomfortable proximity to the main stereo mix paths. Broadcasters face this same problem all the time. All their sources are hostile unless brought up on air.

This is unrelated crosstalk, where the bleeding signal is totally dissimilar and irrelevant to the interfered signal. Basically, if any unrelated crosstalk is audible above system background noise, it will be noticed.

A fairly recent and insidious sort of unrelated crosstalk comes in the forms of assorted chirps, buzzes, and sizzles stemming from the relentless march of digits into console design and operations. The Society of Motion Picture and Television Engineers (SMPTE) time codes and automation codes were bad enough, but trying to get computer clock droning and *vdu* squeaks out of the mixing buses and audio paths is not one of life's most enjoyable tasks.

Designing it out in the first place is the only way to deal with computer noise:

1. Make sure all the logic grounds and analog grounds interrelation makes sense or are tailed back separately and never meet.
2. Scrutinize printed circuit layouts to make sure there are no digital signals adjacent to or on the direct opposite side of the board to any analog signal.
3. Intersperse lots of ground traces.
4. Screen high-current high-speed digital signals.
5. Try to allow only static digital control lines onto analog boards—this means decoding digital buses elsewhere other than on audio boards.
6. Ground-plane everywhere there is board space.
7. Choose logic families—or at least interface devices—that are low current and devoid of large power-rail gulps. CMOS is just fine.
8. Decouple everything for all signals—decouple digital for AF and analog for RF.
9. Work on your karma.

25.14.5 Quantifying Crosstalk

"If you can hear it or measure it, it's failed." Such is the empirical crosstalk test. A more formal test was originally the test for interchannel crosstalk (i.e., between any channels in a console); it's also used for any dissimilar path crosstalk measurements. In short, it asks for better than 60 dB of isolation of 6 kHz between the paths, measured with a standard peak program meter (PPM) with a CCIR 468 weighting filter in line. Since this CCIR curve has 12 dB of gain at its crest (at 6 kHz), the specification is actually calling for better than 72 dB of isolation at 6 kHz, which is neither easy nor often realistic. Such a figure is occasionally not far above system noise floors. Remember, it's a peak measurement; an rms measurement would be some 7–10 dB lower. Nobody said it was going to be easy. Crosstalk's a tough problem.

25.14.6 Meters

Some indication to the operator of the signal levels running through the console and, most importantly, the levels that are being sent to other places is necessary. In Fig. 25-92 a pair of level meter feeds are taken from the top of the dim switches; thus, they follow monitoring. A further pair permanently hung across the main stereo mix output is optional. It's customary to provide metering facilities on each channel; in this design the feed is taken following the monitor path source/return switching. This allows level indication of what is going to a tape track during recording and an "all is well" playback display. Gazing at a row of meters hanging off a multitrack playback, it's surprisingly easy to tell what each is indicating. This is an important cue to a mixing engineer.

There are two basic types of meters, both evolving around the same period on opposing sides of the Atlantic. Each tells the observer entirely different things. Nearly every other sort of audio level indicator (LCD displays, rows of LEDs etc.) nods in style toward one or other of these, (see Chapter 26).

25.14.6.1 VU Meters

Volume unit (VU) meters evolved as a standard in the United States by Bell Telephone Laboratories. A need was shown for a consistent instrument to measure audio levels on lines; it is pictured in Fig. 25-93. The VU meter has a quite tightly defined specification, even down to the buff color of the scale! It is the ubiquitous style of meter that finds itself everywhere, from broadcast consoles to cassette players, but with very few of the interpretations actually bearing much resemblance to the original Bell Laboratories' intentions. It might only consist of a light termination, a rectifier, and a moving-coil meter, but at least the characteristics of all were well defined; enough to be called a standard. Its inception was in Ma Bell's self-defense; she needed some sort of consistency to the levels being hurled down her lines.

In essence, it was a meter only valid for hanging across 600 Ω transmission lines; the 0 VU marking indicates an actual line power level of +4 dBm. The attack and decay times (the time taken for the meter to indicate a steady input signal of 0 VU accurately and the time taken for the needle to fall back afterwards) are some 300 ms. This time is based predominantly on the physical meter ballistics and happens to correlate quite nicely with the level-sensing integration time of the human ear. The VU is intended to give an approximation of how subjectively loud different pieces of program material are in order to match them evenly. This it does quite well. What it doesn't do is give any idea of the actual signal level. The relatively leisurely integration time misses most transients altogether with the consequence that a VU meter will underindicate actual signal level; depending on program material this can be by as much as 20 dB (on impulses and transients—snare drums spring to mind), 12–15 dB on piano, and 8–10 dB on spoken voice.

The underread is unimportant in the respect that the VU does allow subjective level matching and is very easy to read and use. On a purely technical level, the rectifier in a neat, unbuffered VU meter hung straight across a 600 Ω line imbues a serious amount of distortion (some 0.3%) to the program material. This has become more and more of an embarrassment over the years. It is less of a problem with zero-impedance feeds, but, unbuffered, it is still evident.

25.14.6.2 Peak Program Meter (PPM)

The peak program meter, Fig. 25-94, was the British Broadcasting Corporation's answer to the same problem—the BS4297 spec. PPM differs from the VU in three very important respects:

1. The PPM is a peak-reading instrument, capable of accurately displaying signal transients. Correspondingly, it has a very short attack time, coupled with a long fallback decay time to give a chance to see the peaks once it has captured them.
2. The PPM is black.
3. The PPM has a logarithmic scale, allowing accurate signal-level measurements to be made over all the scale range.

The scale consists of seven marks, numbered 1 to 7, each division representing 4 dB level change. PPM 4, the middle mark, is set to indicate 0 dBm/0 dBu. The normal operational maximum signal limit is PPM 6, or +8 dBu.

As accurate and as useful as the PPM is, operators have to consult a list of peak levels for different types of program material (e.g., PPM 5 to 5½ for speech and PPM 4 for heavily compressed pop music) in order to perform the same function as a VU meter, which is subjective level matching. A VU meter makes these adjustments automatically since, although it's worthless for peaks, it follows program density, or loudness, well.

Virtually every other level-indicating device emulates either the VU or the PPM characteristics—or both. There are other European meters—peak reading and log scales over a very wide (40 dB) range and with a longer fallback time than the PPM. Using one of these for the first time is unnerving; audience noise and even studio noise often make the meter dither some way off the bottom stop—a very unusual sight to eyes used to VUs and PPMs.

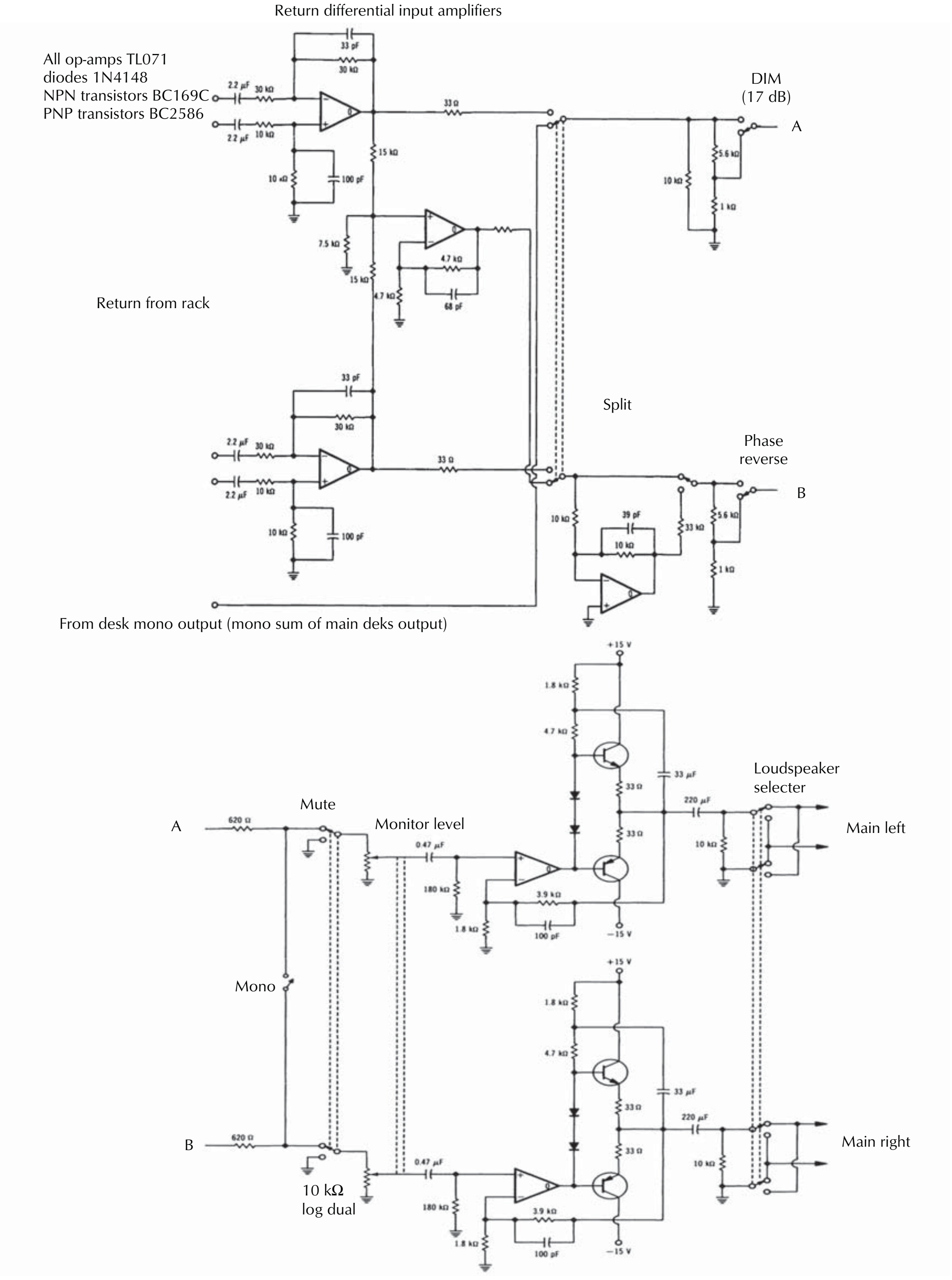

Return differential input amplifiers
All op-amps TL071
diodes 1N4148
NPN transistors BC169C
PNP transistors BC2586
33 pF
30 kΩ
2.2 μF
30 kΩ
10 kΩ
33 Ω
15 kΩ
10 kΩ
100 pF
7.5 kΩ
15 kΩ
4.7 kΩ
4.7 kΩ
68 pF
Return from rack
DIM
(17 dB)
A
10 kΩ
5.6 kΩ
1 kΩ
Split
Phase
reverse
B
39 pF
10 kΩ
33 kΩ
From desk mono output (mono sum of main deks output)
+15 V
1.8 kΩ
4.7 kΩ
33 μF
33 Ω
220 μF
Loudspeaker
selecter
Main left
A
620 Ω
Mute
Monitor level
0.47 μF
180 kΩ
1.8 kΩ
3.9 kΩ
100 pF
−15 V
Mono
B
10 kΩ
log dual
Main right

Figure 25-92. Monitor audio path (desk end).

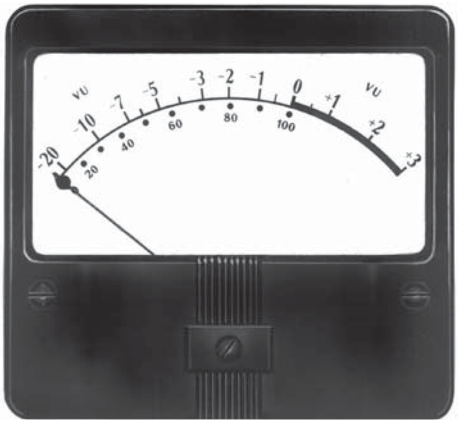


Figure 25-93. Standard VU meter.

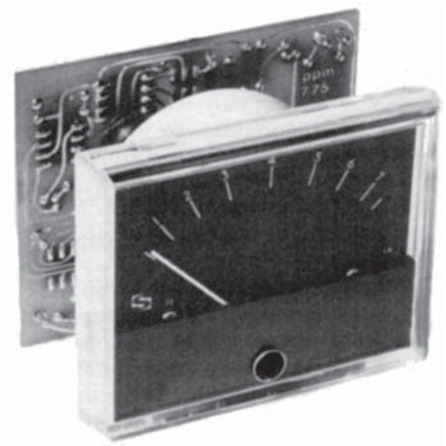


Figure 25-94. Peak program meter.

Proper American broadcasters have taken quite a fancy to a mutant PPM that is similar in dynamic characteristics to BS4297, but with the level for the various marks elevated by 8 dB. The marks give actual level values (up to a maximum of +16 dB whereupon it's painted red) instead of the familiar 1 to 7. This is, it is given to be believed, so that the signal levels generated from control areas using these meters are similar to those from older areas using (curiously nonstandard) +8 dBm referred VU meters. Such are the levels they are used to sending down interstudio and telephone lines.

The elevated-level PPM is an idea with some merit when most of the material dealt with is prerecorded and fairly predictable in level so it does not require an awful lot of head room.

25.14.6.3 Other Metering

A preponderance of LED or FIP bar-graph-style metering, or imitations thereof on GUI (graphical user interface) screens, has caused nearly any sense of adherence to any sort of accepted standard to be abandoned. Most, and especially those on digital equipment, have a very short—if not zero—attack time, which can give rise to misleading readings, and such a wide range of arbitrary release times that it is very difficult to interpret their indications at all. The most an operator can hope to do is to keep it dancing without pinning it. About the best one can rely on is that the top of the meter represents 0 dBfs—or digital full-scale; nominal levels are either −18 dBfs or −20 dBfs (close enough in fader-pushing terms if a world apart technically) depending on whether the influencing force previously used black or beige meters, respectively.

25.15 A Typical Multitrack Console Described

Elsewhere in this chapter, versions of practically every kind of electronic subsystem that finds its way into today's mixing consoles has been described, explored, and analyzed. Here is a description of a complete commercial multitrack mixing console, together with the trials and tribulations of dealing with the electronics as part of an overall system having a life and needs of its own.

A system can be defined as a means of reducing the versatility of its component parts. Ideally, there should be no system, but practicality dictates that there must be one. The thought is mortifying: hundreds of elements, the microphone amplifiers, differential input amplifiers, line amplifiers, equalizers, filters, and routing matrices roaming loose and needing to be coupled together for each individual operational requirement.

We need a saving grace, and fortunately there is one. Engineering and balancing habits are pretty well entrenched, giving rise to a few well-defined, commonly used elemental combinations. Rationalizing these combinations and arranging easy selection of them as necessary is a good compromise. We've not so much lost versatility as gained a family of operating modes.

25.15.1 Channel System Example

This entire channel subsystem relies on the electronic switching elements used being entirely transparent, noiseless, distortionless, clickless, and other impossibilities. Noise due to the potentiometric CMOS switching employed here is very largely due to the individual summing amplifiers, scaled by the gain asked of them.

Noise resulting from them is defined to low (−100 dBu or better) floor levels—fairly meaningless under the stampede of typical front-end or machine noise.

Distortion is primarily due to the automodulation of the CMOS transmission gates; that is, the path resistance varies with the instantaneous signal voltage. This, at zero level, is typically a nonsensical value. Both the harmonic and intermodulation products are almost unmeasurably low principally because of the near virtual-ground operation of the active CMOS elements. There is no voltage swing, no automodulation.

25.15.2 Function Modes

Reference should be made to Figs. 25-95 through 25-98 during this discussion of the channel system. These illustrations show the overall channel in block diagram form and the various ways the circuit blocks are configured for the different functions expected of the channel in use. Fig. 25-95 has all the reconfiguration represented by diagrammatically accurate but forbiddingly incomprehensible mechanical switching. Figs. 25-96 and 25-97 replace those in the main signal paths with electronic switching elements, which may seem more or less of a jungle, dependent on whether you were brought up on hard-gold contacts or silicon.

Certainly there are fewer electronic switchpoints than there were mechanical. This rationalization is primarily due to yet another incursion of esoteric (for audio) digital things.

A simplified representation of the four basic channel operating modes is given in Fig. 25-98A for recording, Fig. 25-98B for mixdown/direct to stereo, and Fig. 25-98C for overdubbing. The Xs show the switching points. Briefly, main multitrack operating modes and their implementation in this system are outlined here.

25.15.3 Recording Mode

In the recording mode, the object is to get a live source (e.g., microphone) through the signal modification chain (i.e., limiting, equalization) and on to a track or tracks of the multitrack machine. Level control on this path is by the main fader (or VCA fader if automation is applicable). Before and after monitoring of the tape track dedicated to the channel is routed onto the main stereo monitoring/mix-bus via the secondary level control.

25.15.4 Mixdown Mode

The machine return is brought through the modification chain and mixed onto the main stereo monitoring/mix-bus via the main/VCA fader. The machine monitoring chain is disabled.

Since a major justification for keeping the multitrack routing open during mixdown is to provide additional effects feeds, this is best served if the secondary level control is fed post main fader and post mute/solo switching. To enable this, a crossfeed electronic routing is included in Fig. 25-98B. However, independent control is restored when required if a fader reverse is called.

Another mode, direct to stereo, is a derivative of mixdown. It enables live sources to be mixed straight on to the main stereo bus, obviating the need to use multitrack routing.

25.15.5 Overdub Mode

A halfway house between record and mixdown, the overdub mode is intended for use when most of the console is in mixdown but individual channels are being laid or touched up. The signal flow is the same as in the record mode, only with the main/VCA and secondary level controls interchanged. The main/VCA fader in this mode, therefore, controls the monitor feed into the main stereo mix-bus, which ties in with the operation of this fader on all the other channels that are in mixdown.

A handy interlock exists in this mode to facilitate single button drop in. When the channel system function is selected to overdub and the monitoring path is set to A check (machine input), a relay closing pair is made that may be plumbed into the remote control access of the machine. Provided the track is armed ready, hitting A check automatically drops the machine into record simultaneously.

25.15.6 Logic Control

This particular console design was intended to be capable of running with or without control by a microprocessor. Much of the localized logic dealing with switches, control, and indication encompasses the necessity to stand alone and in some instances to provide an optional microprocessor with a means of reading back actual console status. Using conventional switch-matrix sensing and latched control/indicator driving by a microcontroller could make redundant much of this hardware, if permanent processor control were envisaged. Similarly, much of the discrete logic

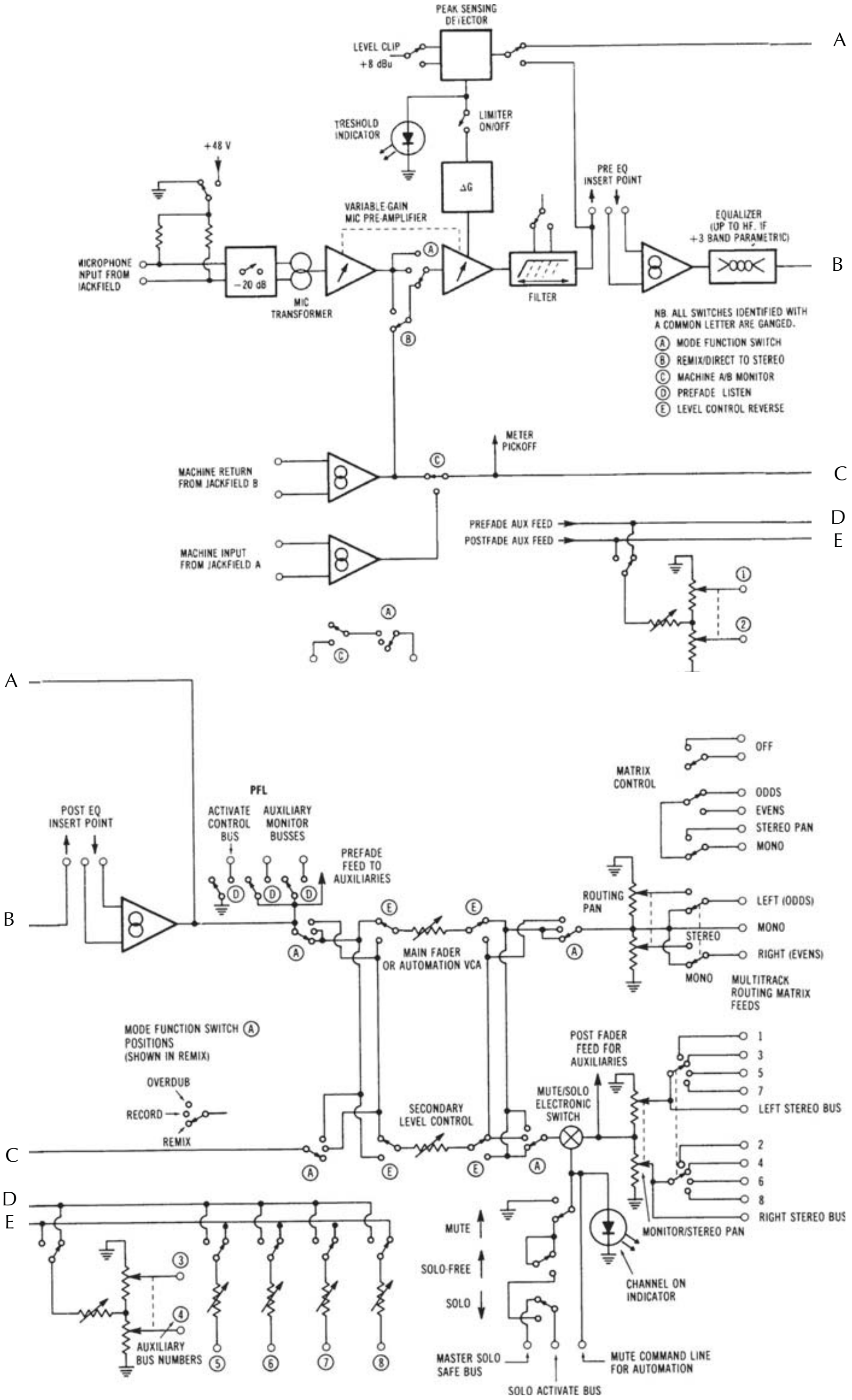

PEAK SENSING DETECTOR
LEVEL CLIP
+8 dBu
A
TRESHOLD INDICATOR
LIMITER ON/OFF
+48 V
ΔG
PRE EQ INSERT POINT
VARIABLE-GAIN MIC PRE-AMPLIFIER
EQUALIZER (UP TO HF. IF +3 BAND PARAMETRIC)
MICROPHONE INPUT FROM JACKFIELD
−20 dB
Ⓐ
B
FILTER
MIC TRANSFORMER
Ⓑ
NB. ALL SWITCHES IDENTIFIED WITH A COMMON LETTER ARE GANGED.
Ⓐ MODE FUNCTION SWITCH
Ⓑ REMIX/DIRECT TO STEREO
Ⓒ MACHINE A/B MONITOR
Ⓓ PREFADE LISTEN
Ⓔ LEVEL CONTROL REVERSE
METER PICKOFF
MACHINE RETURN FROM JACKFIELD B
Ⓒ
C
PREFADE AUX FEED
POSTFADE AUX FEED
D
E
MACHINE INPUT FROM JACKFIELD A
①
②
Ⓐ
Ⓒ
A
OFF
MATRIX CONTROL
ODDS
EVENS
STEREO PAN
MONO
PFL
POST EQ INSERT POINT
ACTIVATE CONTROL BUS
AUXILIARY MONITOR BUSSES
PREFADE FEED TO AUXILIARIES
Ⓓ Ⓓ Ⓓ
B
Ⓔ
Ⓔ
ROUTING PAN
LEFT (ODDS)
MONO
STEREO
RIGHT (EVENS)
Ⓐ
MAIN FADER OR AUTOMATION VCA
Ⓐ
MONO
MULTITRACK ROUTING MATRIX FEEDS
MODE FUNCTION SWITCH Ⓐ POSITIONS (SHOWN IN REMIX)
POST FADER FEED FOR AUXILIARIES
1
3
5
7
LEFT STEREO BUS
OVERDUB
RECORD
REMIX
SECONDARY LEVEL CONTROL
MUTE/SOLO ELECTRONIC SWITCH
C
Ⓐ
Ⓔ
Ⓔ
Ⓐ
2
4
6
8
RIGHT STEREO BUS
D
E
MUTE
MONITOR/STEREO PAN
③
SOLO-FREE
CHANNEL ON INDICATOR
SOLO
④
AUXILIARY BUS NUMBERS
⑤
⑥
⑦
⑧
MASTER SOLO SAFE BUS
MUTE COMMAND LINE FOR AUTOMATION
SOLO ACTIVATE BUS

Figure 25-95. In-line multitrack recording and monitoring channel.

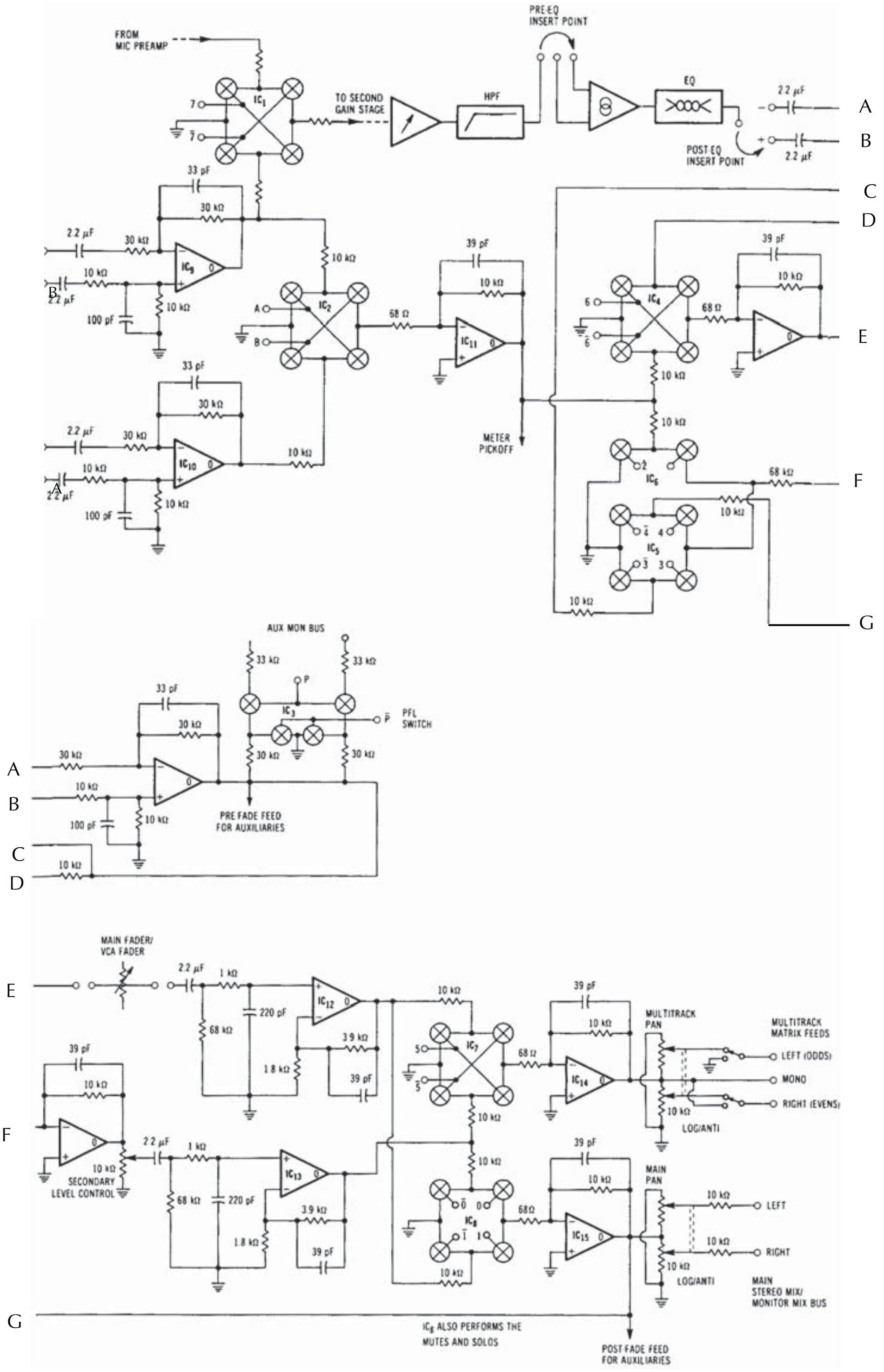
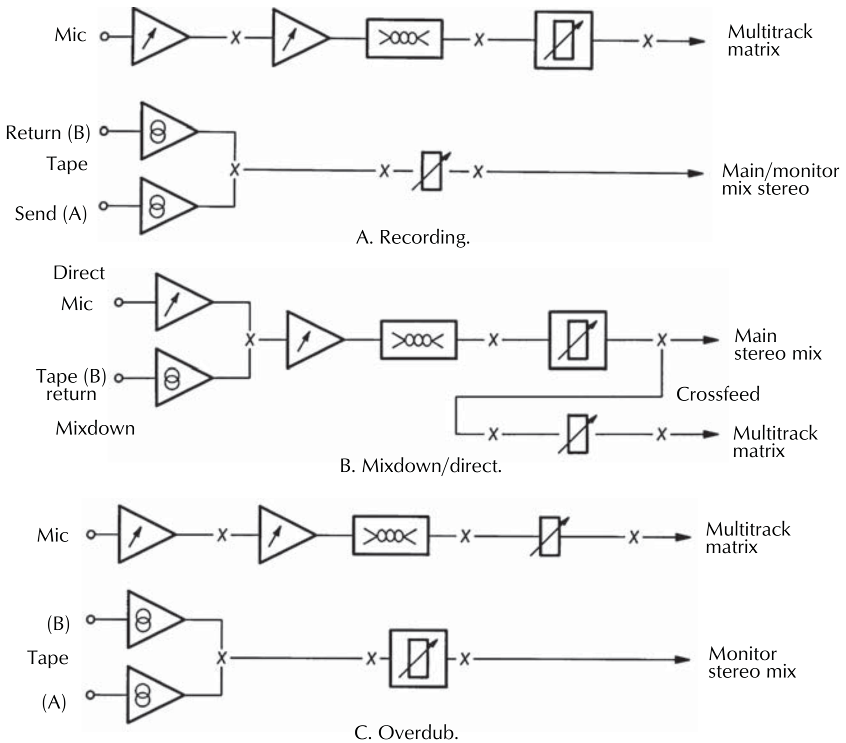

FROM MIC PREAMP
IC1
7
7̄
TO SECOND GAIN STAGE
HPF
PRE-EQ INSERT POINT
EQ
2.2 µF
A
B
POST-EQ INSERT POINT
2.2 µF
33 pF
30 kΩ
2.2 µF
30 kΩ
IC9
10 kΩ
2.2 µF
10 kΩ
100 pF
10 kΩ
IC2
A
B
68 Ω
39 pF
10 kΩ
IC11
C
D
39 pF
10 kΩ
IC4
6
6̄
68 Ω
E
33 pF
30 kΩ
2.2 µF
30 kΩ
IC10
10 kΩ
10 kΩ
2.2 µF
100 pF
10 kΩ
METER PICKOFF
10 kΩ
10 kΩ
IC6
2
68 kΩ
F
10 kΩ
IC5
4̄
4
3̄
3
10 kΩ
G
AUX MON BUS
33 kΩ
33 kΩ
P
33 pF
IC3
P̄
PFL SWITCH
30 kΩ
30 kΩ
30 kΩ
30 kΩ
A
B
10 kΩ
100 pF
10 kΩ
PRE-FADE FEED FOR AUXILIARIES
C
10 kΩ
D
MAIN FADER/ VCA FADER
E
2.2 µF
1 kΩ
IC12
220 pF
68 kΩ
3.9 kΩ
1.8 kΩ
39 pF
10 kΩ
IC7
5
5̄
68 Ω
39 pF
10 kΩ
IC14
MULTITRACK PAN
MULTITRACK MATRIX FEEDS
LEFT (ODDS)
MONO
RIGHT (EVENS)
10 kΩ
LOG/ANTI
39 pF
10 kΩ
F
10 kΩ
SECONDARY LEVEL CONTROL
2.2 µF
1 kΩ
IC13
68 kΩ
220 pF
3.9 kΩ
1.8 kΩ
39 pF
10 kΩ
10 kΩ
IC8
0̄
0
1̄
1
68Ω
39 pF
10 kΩ
IC15
MAIN PAN
10 kΩ
LEFT
10 kΩ
RIGHT
10 kΩ
LOG/ANTI
MAIN STEREO MIX/ MONITOR MIX BUS
10 kΩ
G
IC8 ALSO PERFORMS THE MUTES AND SOLOS
POST-FADE FEED FOR AUXILIARIES

Figure 25-96. Channel system—audio paths.



A. Recording.

B. Mixdown/direct.

C. Overdub.

Figure 25-97. Channel function modes.

elements would today be wrapped up in programmable CPLD and FPGA (field-programmable gate array) packages for flexibility and to minimize part count and board space. The discrete logic is left detailed here, however, on the basis that it would be useful to have an idea of what sort of logic would need to be programmed into these parts!

A distinction is made in Figs. 25-96 and 25-97 between the analog signal switches and their digital control electronics not purely because of the differing disciplines but for clarity's sake; that is, to avoid too many lines running all over the place on the drawing.

Each top-panel switch is a momentary-action touch switch with an associated LED indicator (with the exception of the function mode switch). The toggle push-on and push-off characteristic is provided by the basic debouncer/flip-flop circuit, as shown in Fig. 25-99. This action is not only fun, play-worthy, and therefore, fashionable, it also scores in a couple of other important respects:

Cost. The combination of a small, mechanically simple, nonlatching, push-to-make switch and a fairly small number of silicon bits is much less expensive than latching pushbutton switches.

Versatility. Using electronic latching rather than mechanical catches makes remote/automatic function presetting and triggering comparatively simple.

25.15.7 Switch Debouncing

Debouncing is removing the ragged edges from a switching signal. Switch contacts do not simply make contact when pressed and break contact on release. The two bits of metal grind against each other or bounce a few times while moving together or apart, resulting in a series of ragged, spiky "almost contacts" rather than simply touch or not touch.

Ordinarily, this doesn't matter too much, but, if the switch is feeding a bistable flip-flop (as here), the fun begins. Flip-flops are usually edge triggers; on a positive-going transition, another pulse flops it back and so on. A string of rapid, unpredictable pulses, as provided by nearly any mechanical switch, sends flip-flops frantic.

Retarding the switch with time constants is nearly foolproof, but the arrangement in Fig. 25-99 is practically faultless. The 4098 contains two monostables, which are handy since the 4013 contains two flip-flops. It can sense either positive or negative transitions, posi-

if

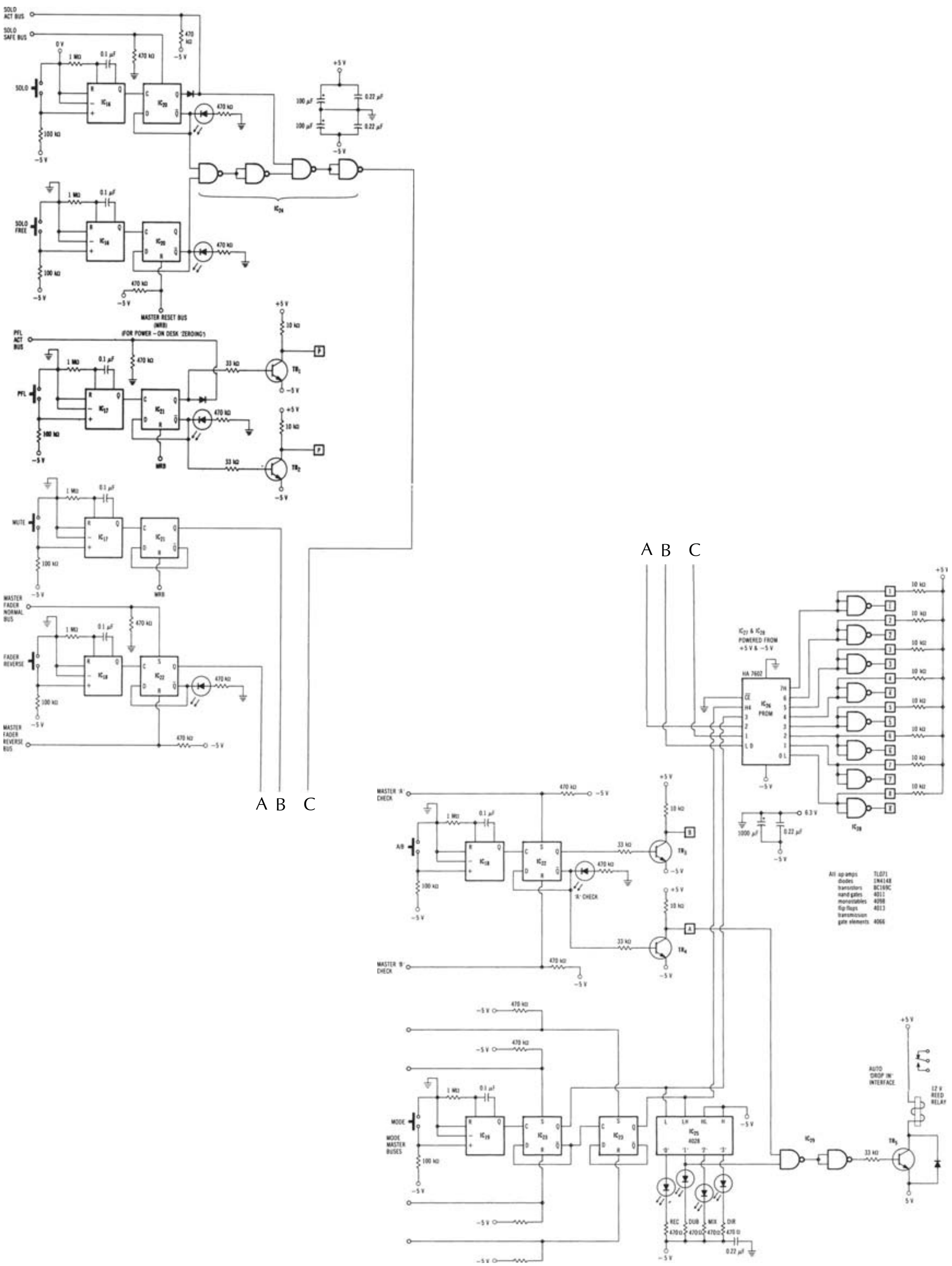

SOLO ACT BUS
SOLO SAFE BUS
0 V
1 MΩ
0.1 µF
470 kΩ
470 kΩ
–5 V
SOLO
IC18
IC20
100 kΩ
–5 V
+5 V
100 µF
0.22 µF
IC24
SOLO FREE
MASTER RESET BUS (MRB) (FOR POWER – ON DESK 'ZEROING')
PFL ACT BUS
PFL
IC17
IC21
33 kΩ
10 kΩ
TR1
TR2
P
MRB
MUTE
MASTER FADER NORMAL BUS
FADER REVERSE
IC22
MASTER FADER REVERSE BUS
A B C
A B C
IC27 & IC28 POWERED FROM +5 V & –5 V
HA 7602
IC26 PROM
6.3 V
1000 µF
0.22 µF
IC28
MASTER 'A' CHECK
A/B
TR3
TR4
'A' CHECK
MASTER 'B' CHECK
MODE
MODE MASTER BUSES
IC19
IC23
IC25 4028
REC
DUB
MIX
DIR
IC29
33 kΩ
TR5
AUTO 'DROP IN' INTERFACE
12 V REED RELAY
5 V
All op amps TL071
diodes 1N4148
transistors BC169C
nand gates 4011
monostables 4098
flip flops 4013
transmission gate elements 4066

Figure 25-98. Channel system—control logic.

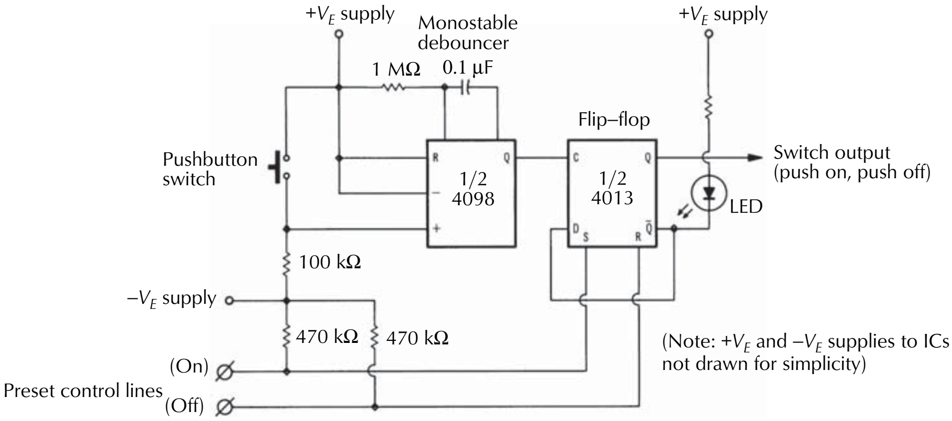


Figure 25-99. Push button interface circuit.

tive in this application, catch the very first input transition, and stuff out a uniform, clean, predictable clock pulse for the flip-flop. Subsequent bounces merely extend the output pulse slightly but don't generate any spurious output transitions. An alternative would be the interspersing of Schmitt trigger buffers between the switches and the flip-flops. These have a very wide hysteresis, which in conjunction with some R/C retarding can also provide surprise-free toggling.

Flip-flops can have their outputs jammed by stuffing the required state-up set (making the *Q* output go positive) or reset (negative)—an invitation for remote processor control.

25.15.8 Logic Sense

Some of the logic in this particular design is unconventional, all done in the name of reducing component count, largely obviating level-shifting transistors while maintaining the inviolable ground-for-active law of control interfacing. This is a common-sense rule that simply means that any accessible control line should just need to be taken to some reasonable ground in order to activate whatever it's supposed to—not to a specific voltage above or below ground. This helps avoid the "should this go to +5 V or –24 V" routine, while greatly simplifying system design—grounds are omnipresent.

The main reason for the unusual logic powering, Fig. 25-97, stems from the use of a bipolar PROM in the assignment logic. This needs a tightly controlled 5 V supply, unlike CMOS ICs, which will run off nearly anything with volts on it.

25.15.9 What Is a PROM?

PROMs (or programmable read-only memories) are digital devices used extensively in computer technology for storing individual items of information or sequences of information that are regularly referred to:

- *Memory* is self-explanatory.
- *Read-only* means that in normal operation it's only possible to retrieve the information that's stored. New information cannot be put in or the contents modified.
- *Programmable* means, given the right gear and software, prepared information can be written into the PROM. The type used in this design can't be restuffed though, since the programming is achieved by literally blowing tiny internal fuses in the shape of the data. This seeming inversatility is reasonable with such devices where the device cost is cheap compared with programming costs (human time).

The information stored is, of course, binary in nature—a 0 or a 1, up or down, there or not, and so on. The number of these binary bits contained in each PROM can be in the millions. Four megabit proms are now common. For this channel system control, the PROM used stores 256 bits, which, in fact, is still a bit of overkill, but they don't really come much smaller.

This baby PROM, a Harris 7602, is much like most adult PROMs in that the bits are organized internally in chunks eight wide, as in a digital word (byte). Eight happens to be the byte width of most popular microcontrollers. In the baby PROM there are 32 such bytes of stored data (32 × 8 = 256), each being accessible with a specific 5-bit-wide address code (given by the binary numbers from 0 to 31). For any of up to 32 command states, preprogrammed responses for eight output lines are immediately accessible.

This particular type of baby PROM is usually used at the top end of microprocessor memory maps where a page (256 bytes) is given over to the function of the processor vectors, such as interrupts. As an example, if the processor receives a *nonmaskable interrupt* (NMI), it usually means "Panic! The power is collapsing!" or some other similar situation. NMI makes the processor look at a certain address in the page of the baby PROM, which tells it where to find in memory a program to save the environment (i.e., hide safely all the crucial operating data, quickly).

In the context of this channel system, the PROM outputs drive the analog switches (organized per Fig. 25-96) to route and control the channel and monitor signal paths through the system elements. This occurs in accordance with and under the command of the PROM address inputs, which are indicators of selected channel function (record/mixdown/overdub), local or remote fader reverse commands, and, importantly, mute and solo status.

Most of the control logic is still done in hardware, largely consisting of jammable debouncer/flip-flops. For the channel function control, a single pushbutton that steps through the four functions is realized by a simple 2 bit counter (IC23 in Fig. 25-97). This generates a 2 bit code that feeds both the PROM control inputs and a 4028 binary to decimal decoder IC25, which drives the relative status indicating front-panel LEDs.

Solo, solo-free, and solo-safe are dealt with in IC16, IC20, and IC24, but the relevant action on the analog circuitry is still executed via the PROM. It can be deduced that the solo command and mute of the PROM do just the same thing, resulting in a fair number of duplicated and redundant program codes within the PROM. At least this gives room for expansion or function modification (if and when required) by simple card link changes and a differently programmed PROM. Here is one of digital's great strengths—the future capability of "chameleoning" a system simply by software changes, not by hardware: a built-in upgrade path.

25.15.10 Logic Meets Analog

The 7602 PROM hangs between logic ground and −5 V (of the split ±5 V logic supply), thus necessitating all input feeds to be similar in swing—0 to −5 V. All the drive logic flip-flops, debouncer, and master bus logic are similarly powered.

Analog transmission gates, such as the design of Fig. 25-96, are required to pass (and stop) analog signals referred to ground and, therefore, of both polarities, so the gates have to be fed from a split supply (in this instance, the ±5 V logic supply).

Converting between the 0 and −5 V logic and the ±5 V control voltage swing needed by the gates is done by using the open-collector output drives of the PROM, Fig. 25-100. Open-collector is exactly that—there is no positive output pull-up internal to this PROM. The idea is that it may be paralleled with other open-collector devices in a wired-OR bus configuration. When the output transistor is turned off, the collector is at a high-impedance state. The collector is pulled up an extra 5 V above the internal supply of the PROM. When the transistor turns on, the collector dutifully zaps down to the −5 V supply. It doesn't care what is at the other end of the load pulling resistor provided it isn't of excessive potential (12 V is safe; the output ports are, in fact, the programming path with these devices and much above that may induce some involuntary reprogramming).

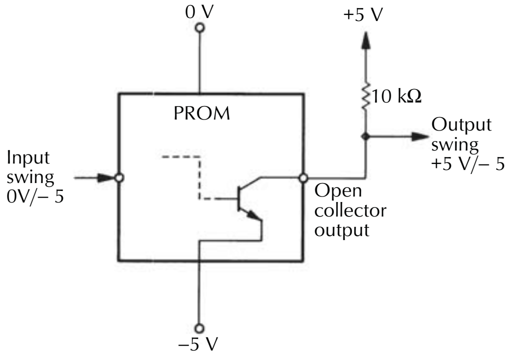


Figure 25-100. Input/output termination (unipolar to bipolar control swing conversion utilizing PROM open- collector output).

Some of the analog switches are driven directly off the PROM outputs, while others have the necessary inverse-switching feed provided by conventional inverters.

As a note to the unwary, bipolar memories such as the 7602 use a lot of power when being switched. This explains the large amount of decoupling festooned around it and the logic supply generally. Needless to say, the analog transmission gates are referred to audio ground, not the click-infested logic ground, despite the fact that they are powered off the logic supply.

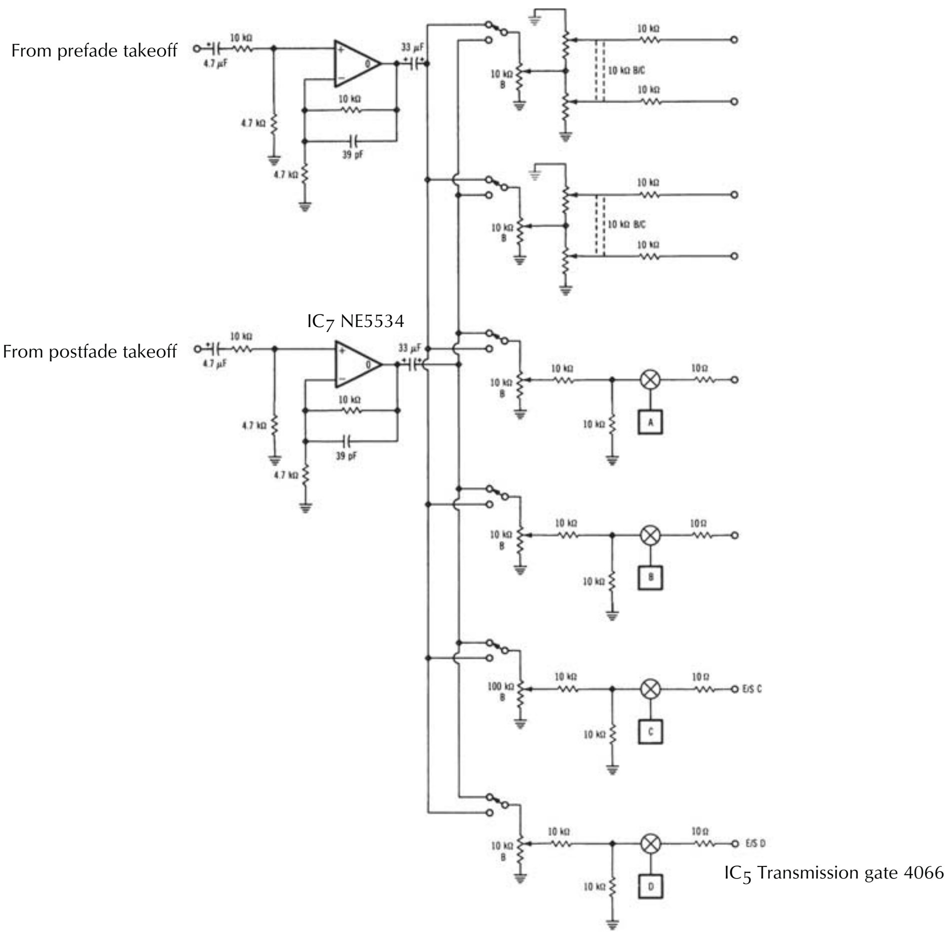

From prefade takeoff
4.7 µF
10 kΩ
4.7 kΩ
10 kΩ
39 pF
4.7 kΩ
33 µF
10 kΩ B
10 kΩ
10 kΩ B/C
10 kΩ
10 kΩ B
10 kΩ
10 kΩ B/C
10 kΩ
IC7 NE5534
From postfade takeoff
4.7 µF
10 kΩ
4.7 kΩ
10 kΩ
39 pF
4.7 kΩ
33 µF
10 kΩ B
10 kΩ
10 kΩ
10Ω
A
10 kΩ B
10 kΩ
10 kΩ
10Ω
B
100 kΩ B
10 kΩ
10 kΩ
10Ω
E/S C
C
10 kΩ B
10 kΩ
10 kΩ
10Ω
E/S D
D
IC5 Transmission gate 4066

Figure 25-101. Channel auxiliary sends and logic control—Auxiliary sends channel.

25.15.11 Auxiliary Channel Feeds

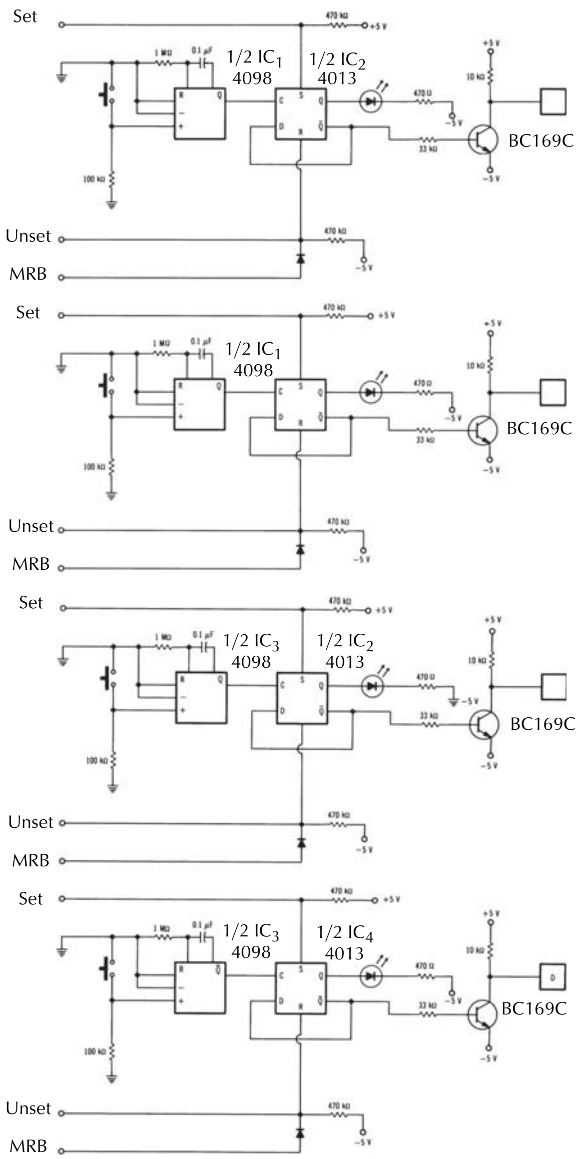
Two prefade (and so premute) feeds are provided on each channel, each with a level control and pannable across a stereo pair of mix-buses. This provides a versatile facility enabling separate stereo foldbacks or four separate feeds. Each of the pairs is selectable to postfade should extra effect feeds be needed during a heavy mixdown, whereupon they will also be subject to channel mutes. A few effects such as stereo reverberation plate/black box would benefit enormously in operation if they could be sourced from stereo auxiliary feeds such as these.

Four individual postfade effect feeds are individually mutable (locally or remotely), individually level controlled, and selectable to prefade.

Effect feeds are quite often switched during mixes; consequently, analog transmission gates are used, facilitating automation. Local activation is achieved through the debounce/latch arrangement used extensively in the channel mode switching, Figs. 25-99 and 25-101. The latch output drives a simple, single-element transmission gate. Isolation, crosstalk, and noise criteria are not particularly critical on these feeds, but they still come out quite creditably. The console switch-on master reset bus (MRB) cancels all these feeds, leaving a clean slate rather than the alternative unpredictable hordes of ons, offs, and maybes in the event of a power interruption or control zeroing.

25.15.12 Summing Modules

Much of the actual mixing within the system described so far is self-contained. Multitrack routing, when achieved via a matrix, allows multiple sourcing to any


Set
1 MΩ
0.1 µF
1/2 IC1
4098
1/2 IC2
4013
470 kΩ
+5 V
470 Ω
−5 V
10 kΩ
33 kΩ
BC169C
100 kΩ
Unset
MRB
Set
1/2 IC1
4098
BC169C
Unset
MRB
Set
1/2 IC3
4098
1/2 IC2
4013
BC169C
Unset
MRB
Set
1/2 IC3
4098
1/2 IC4
4013
BC169C
Unset
MRB

B. Effect send logic control.

Figure 25-101. Channel auxiliary sends and logic control—Effect send logic control. (Continued)

chosen group or machine track. A stereo mixdown of all the channels is possible with this method by selecting them to an arbitrary pair of tracks across which the mastering machine is hung. This is, in fact, the mixdown technique used in many console systems whether in-line or discrete monitoring. Although entirely feasible, it is not the manner in which this particular system is intended to be used, and is becoming rarer in the commercial world.

Stereo mixdown is achieved in the same buses as the multitrack monitor mix, the solo monitor function making its home here, too.

A master group module contains the mix-amps, fader, and line amps pertaining to the stereo bus together with sundry other related things, like mono summing (required for a monitor feed) and clean auxiliary bus access for extending the monitor mix (for effect returns or temporary extra channels).

25.15.13 Mixes to Outputs

The virtual-earth mixing buses of the console all end up in identical mix-amp, attenuator, and line amp configurations. The exceptions are the mono sources (effect sends) that have individual master level controls rather than ganged stereo attenuators and the PFL (which does not need a level control because it is a purely monitoring function). These back-end stages are homed in two of the very few on-off system blocks in this design: stereo monitor/mix (with the master fader) and the PFL summing occupy the master module, while the remaining auxiliary functions are summed in the auxiliary master module, Figs. 25-86 and 25-87.

The outputs are taken to the jackfield, where they are normalized to their appropriate destinations and directly bridged by the differential inputs of the monitor selector switching matrix (adjacent to the field). It is assumed that output transformers or electronic balanced output line amps will be implemented.

25.15.14 Master Functions

Fig. 25-102 shows a simple console master function circuitry. All the clever switching is done in the channels, allowing this unit to be little more than switch contacts. No debouncing is necessary since the master buses directly actuate the set and reset latch functions of the channel function registers.

Lockouts are arranged on the fader main/reverse selection and master monitor A and B switching to prevent both of the relevant control buses from being switched at the same time; this could otherwise lead to

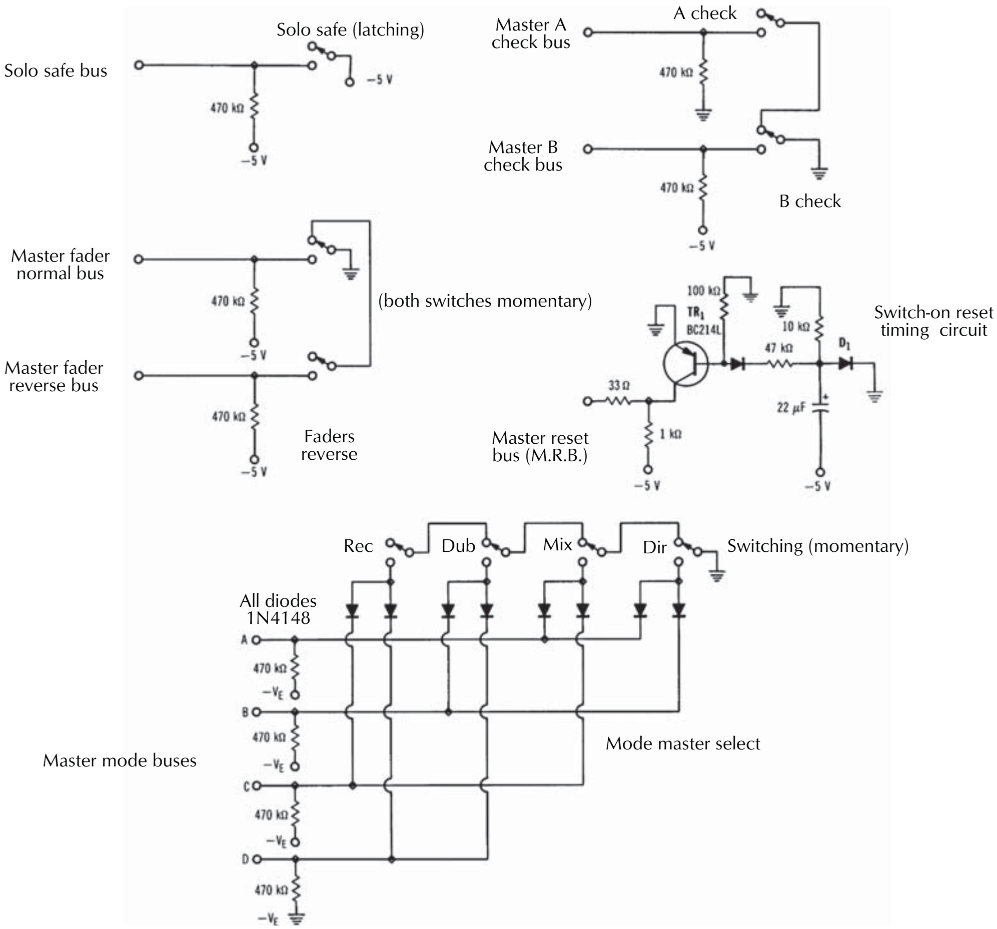

Solo safe (latching)
Solo safe bus
470 kΩ
−5 V
−5 V
Master A check bus
A check
470 kΩ
Master B check bus
470 kΩ
B check
−5 V
Master fader normal bus
470 kΩ
−5 V
(both switches momentary)
Master fader reverse bus
470 kΩ
−5 V
Faders reverse
100 kΩ
TR1
BC214L
10 kΩ
47 kΩ
D1
Switch-on reset timing circuit
33 Ω
22 μF
1 kΩ
Master reset bus (M.R.B.)
−5 V
−5 V
Rec
Dub
Mix
Dir
Switching (momentary)
All diodes 1N4148
A
470 kΩ
−VE
B
470 kΩ
−VE
C
470 kΩ
−VE
D
470 kΩ
−VE
Master mode buses
Mode master select

Figure 25-102. Desk master function control circuitry.

some very odd things happening inside the channel signal routing. Similarly, a ground follow-through lockout arrangement is used on the master function mode selection. Otherwise, the consequences of more than one button being pushed simultaneously would be to select a virtually random mode.

Note that all the switching is to ground from the logic −5 V supply. This interfaces with the majority of the channel logic as described. An important feature is the master reset bus and its control. Ordinarily, an array of random logic circuitry dependent on flip-flops and latches (of which this design is an example) would, on power up, tend to settle into whatever state these registers felt like at the time. The result would depend on device symmetry, temperature, and humidity, but worse still, the results are not usually repeatable. An intriguing exception to this is the knack of CMOS flip-flops to come back up in their previous state after a short power disablement, probably a function of small charge storage.

25.15.15 Power on Reset

Wisdom and common sense dictate that on power up the console should come on neutral, with all channels muted and with monitoring functions such as PFL and solo disabled. The one exception to this is if the console control surface is totally under the control of a computer (all functions), in which case it may be arranged to come up in exactly the state in which it was last turned off, or otherwise lost power. Even in this case, it would be wise to bring the monitoring up muted. As well as providing a frame of reference from which to start reusing the console, it saves all the aggravation of finding the one function that's killing the monitoring. There are few things scarier or more frustrating than a large console that is mysteriously and totally silent. Console mode and basic monitoring conditions can be set up just by pushing the relevant master controls. Processor control of course would afford the console being reinstated to its last used configuration.

TR_1 in Fig. 25-102 grounds the master reset bus (M.R.B.) for as long as the 22 μF capacitor takes to charge up—around a quarter of a second. This charging takes place when the −5 V logic supply appears. Should the supply collapse, the capacitor is rapidly discharged via D_1 ready to reinitialize the M.R.B. signal as soon as power is reestablished.

Although it would be extremely simple to do, no top-panel master reset control is made available because sooner or later someone would hit that button at exactly the wrong moment. This follows the same philosophy that frowns upon a top-panel ac line switch.

25.15.16 Meters and Head room

There are plenty of proprietary meters of the popular standards and types, plus quite a few strange ones, too. It's all a matter of personal preference and the information hopefully gleaned from the assorted needles, lights, and cathode rays dancing before the eyes.

Without jumping into the argument of average versus peak-reading instruments, it is relevant to state that the choice will directly affect the operational levels, the level architecture, the machine lineups, and the various tweaks, notably the input stage limiter threshold in this design. Out of habit, this console was designed with standard PPMs in mind, where the peak operational level throughout the system is expected to be PPM 6, or +8 Bu. Lineup level (i.e., the system and output level for which the front-end gain stage is calibrated) is 0 dBu, PPM 4. This will suit any current or expected PPMs.

VUs are very good for giving an idea of subjective loudness and not worrying you about transients that can often be anything up to 20 dB above the indicated value.

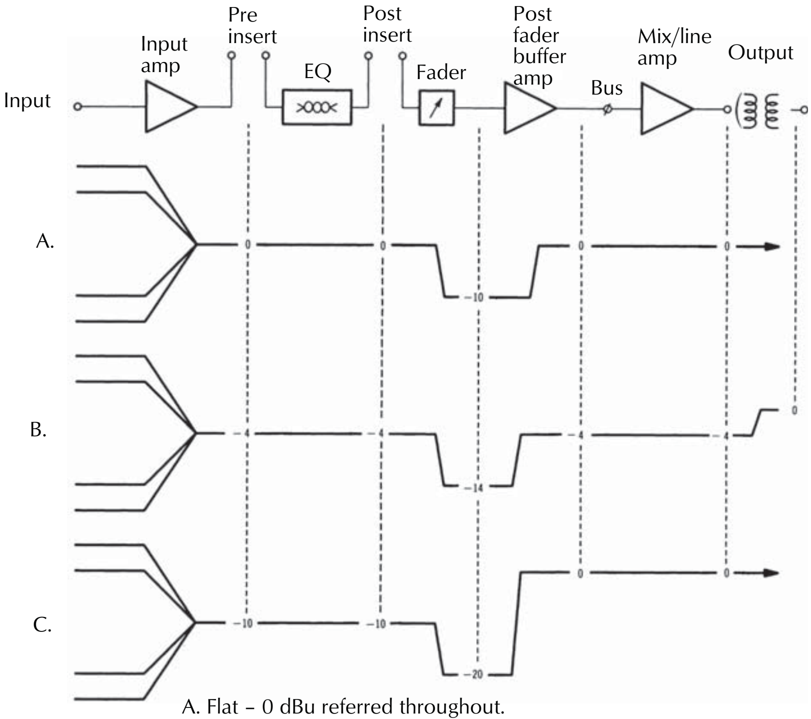
25.15.17 System Level Architecture

Nonunity-level architectures are regrettably necessary under some conditions—detailed here are ways (quite typical fixes within most console designs) that are directly applicable to this described console.

Given standard +4 dBm referred VU meters, under normal operational circumstances, head room in any console is perilously skinny. Various ways of dealing with potentially inadequate head room are in use, Fig. 25-103. A favorite is to run the entire console system at a depressed level, usually −4 dB, the necessary 4 dB makeup at the end being done passively by an output transformer ratio stepup. This is a poor choice for two reasons. The transformer stepup arrangement is overly critical to termination impedance, and the frequency response could suffer with a heavily reactive load such as a long line.

A more modern solution on a similar theme is to adopt a depressed level of −6 dBu and make up the level at the output in a quasi-balanced electronic output stage—in this way head room is not compromised at any point along the way.

Head room is mostly a problem in input channels, before the channel gain-controlling element, the fader. Both ragged unpredictable input sources and equalizer gain gobble up the nonmargin. Hopefully, beyond that point the levels and, hence, the mix are easily and well regulated by the faders. Dropping the channel operating



A. Flat – 0 dBu referred throughout.
B. Depressed – 4 dBu throughout, level made up passively by output transformer.
C. Channel Depressed – 10 dBu in channel, 0 dBu output.

Figure 25-103. System level architecture.

level by 6 dB or 10 dB helps matters tremendously, and the gain is made up either in the mix-amps or the post-fader buffer amps (the latter being normal). This does compromise bus noise (quiescent console output noise), but since the main justification for doing it is the high level of signals present, the pluses outweigh the minuses. This depressed channel system is worthwhile in any circumstance, regardless of metering type, where there is likely to be a great unknown lurking on the end of an input line.

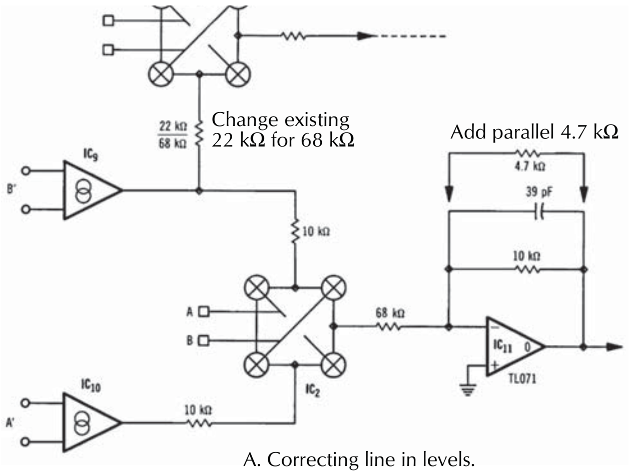
Some of the disadvantages are that all the channel insert points operate at the depressed (10 dBu) level, which may or may not give problems in some less than versatile outboard devices. The more immediate concern is that other internal channel circuits will need adjusting.

Machine line-in feeds from the A′ and B′ input differential amplifiers will need to be dropped by 10 dB. This drop is easily accomplished by altering the values of the resistors around electronic switches to scale down a factor of 3.16 (10 dB), as shown by Fig. 25-104. The PFL bus mix-amp gains are required to increase 10 dB (the extra bus noise here is no great crime), and an extra 10 dB of gain is put into the prefader auxiliary feed buffer amplifiers. Reestablishing main path gain to unity is simply achieved by upping the gain of the post-fader buffer amplifier in Fig. 25-104 and by changing the feedback bottom leg resistors in Fig. 25-96 from 1.8 kΩ to 430 Ω. This provides for 10 dB of fader back-off and the necessary 10 dB reinstatement.

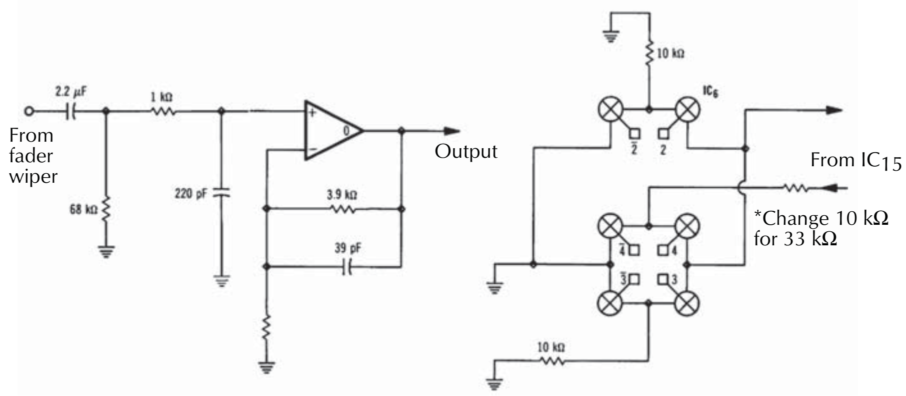
If all that sounds complicated, just bear in mind that it's achieved with gain changes—in this case just with resistor changes. It doesn't matter that the machine monitor differential input amplifiers are still operating at a normal undepressed level. The A′ check is directly monitoring a console output, which is at normal level anyway, so there is no head room problem. As for the B′ check, if we have more level coming back from the machine than we're putting in (A' check), then it's time for realignment.

It is entirely possible to recalculate the values around the differential amplifier to drop 10 dB and still maintain input balance, but that would greatly increase the number of component changes necessary to alter channel system level. This is no mean consideration should you choose to do so on a console of 32 or 48 channels.

Ultimately, it is up to the designer to make the product—this mixer—as transparent and free for the



A. Correcting line in levels.



B. Clippling gain of post fader buffer amplifier.

C. Maintaining correct cross–feed level.

Figure 25-104. Component changes to operate channel at a depressed (–10 dBu) level. (Refer to Fig. 25-96.)

operators as possible. Messing around with system head rooms does somewhat fall into the category of kludge, but as a means to the end of facilitating the music-making process less painful, who can argue? A console is after all a creative tool, not a museum of technical operating standards.

25.16 Consoles and Computers

Varying levels of logical control, remote control, automation, and data storage and recall were common on analog consoles long before digital consoles came along, on which of course there is no option. The "Why?" of such digital control of analog was a much harder question to answer at the time, but decent solutions to most of the requirements were found.

Automated fader systems led to the quest for storing and recalling all console functions, since the instantaneous storage, recall, and automation of console settings had immediate application in several spheres of activity. But by the time this was all achievable with a comparable level of performance as from a nonautomated console, digital mixers were becoming a reality. Since they were of necessity completely programmable, and the control hardware and software was by its very nature already there and in place to do automation to some degree or other, it was a done deal.

What has emerged out of the growth of digital control is the fascinating question of control-surface ergonomics; far from being shackled to the hardware beneath them, now the surfaces can actually be designed to best suit their purpose. Seconds aside, back to back, ten paces then turn, please.

A further set of considerations comes from the wisdom, necessity, and/or desirability of siting the guts of the console (the signal-processing bit) remote from the control surface. Apart from the need for extensive communication between the two (usually attacked by networking-think) the effect on the design of the console architecture is actually surprisingly minor, and the impacts such as they are will be dealt with piecemeal as required. No, the control surface is the real battleground.

25.16.1 Fader Automation

The first victims of automation were the faders. Once heavy multitrack (16/24 track) had become commonplace, a severely limiting factor of human physiology—only ten fingers—proved something of an obstacle in a mixdown situation demanding considerably in excess of that number. The hitherto classic solution—reduction mixes of subgroups of tracks to a more manageable quantity—forces another tape generation; this is not a good idea considering one of multitrack's touted advantages is freedom from bouncing.

To be able to remember, and subsequently modify if need be, fader movements during a mix seemed like a good idea. There were, and still are, two fundamental approaches to this requirement:

1. Remember the physical position of the fader and on recall arrange for it to move physically to its required position.

   This first technique was introduced initially by one major manufacturer (Neve's NECAM system) and with the availability of reasonably economical motorized faders is now fairly widespread. Most others fall broadly into the second camp. Moving fader systems are dearly loved by their users because of their unequivocable indication at all times—by the actual fader positions—of what the system is actually doing. It has one other major benefit—the involuntary hysterical laughter it spontaneously generates from anyone who for the first time sees a swath of motor-driven faders dancing about on their own.

   With the ready availability of such moving faders at cost points suited to nearly every level of application, moving fader automation systems have become the de facto standard for both automated analog and digital mixers.
2. Drive a voltage-controlled amplifier (VCA) from the fader and on recall reapply the appropriate control voltage to the VCA—the fader itself is not then controlling the VCA, Fig. 25-105.

   VCA systems remain viable in analog consoles, though, since they offer advantages at that crucial fader point that moving faders cannot alone fulfill. Although VCA automation systems were once implemented in a purely analog fashion, the fader position values being stored by a PWM or voltage-to-frequency conversion methodology on an analog tape track, these techniques mercifully gave way to digital manipulation and storage as soon as it was practicable.

   A nulling indicator, as described later, is usually employed to match actual VCA gains to that notionally indicated by the fader.

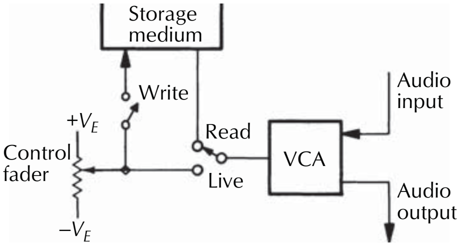


Figure 25-105. Simplistic VCA-type fader automation.

25.16.2 VCAs

Several functions in mixing consoles cry out for a perfect and consistent controllable gain block. In addition to automated fader systems, dynamics control and other analog-controlled gain stages could all benefit by something that looks like Fig. 25-106. It is a black box to which audio is applied, from which audio is extracted, and a control port that determines how much audio is passed. Ideally the law of the control signal should be predictable and consistent. No biasing, no tweaks, no singing, no dancing. Should be easy, right?

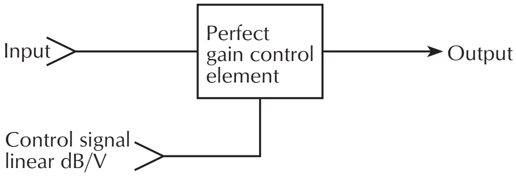
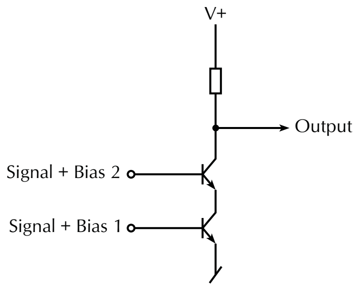


Figure 25-106. Ideal gain control

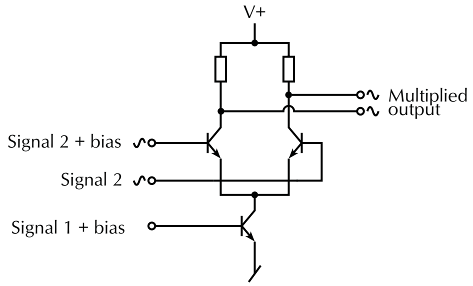
As seen elsewhere, raw active electronic devices can be used as gain-variable stages with varying degrees of success, compromises, and weirdnesses; their limitations are various but notably include limited audio signal handling capability, high distortion, and often nonlinear (or nonsensible) control-voltage laws. In feedback-style automatic gain-reduction circuits such as compressors and limiters, the law of nonlinearity tends to disappear within the servo-loop feedback and have either negligible or—better yet—an interesting effect on the behavior of the circuit in response to stimulation. Effectively biased—to avoid their turned-off regions—FETs can have a square law response, transistors an exponential (logarithmic) response of collector current with respect to applied base voltage. Logarithmic?—dBs are logarithmic!

25.16.2.1 Transistor Junctions

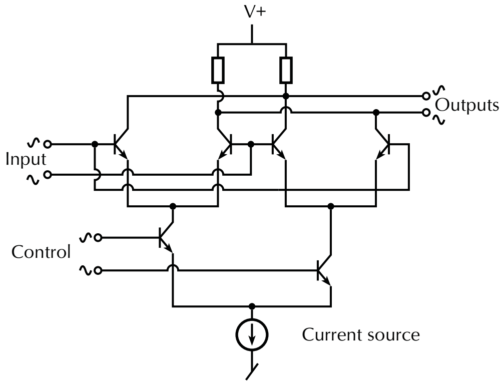
The departure point for the journey to VCA-dom is Fig. 25-107A, which is for our immediate purpose actually pretty useless but elsewhere is known as a *cascode amplifier*. The upper transistor's emitter serves as a nonvoltage varying load to the lower transistor, allowing it to achieve large bandwidth current gain free of Miller effect; the upper transistor (as essentially a common-base amplifier) has no current gain but serves to buffer the load in its collector from the lower transistor, which is busy doing all the work. Varying the base voltage of the upper transistor has little effect on anything other than altering the maximum voltage swing capability on the load, certainly not gain, which is all very much different from the long-tailed pair of Fig. 25-107B. Note that the upper stage is and can be used differentially, as is the output, but it works single-ended too. Here the current through the load is modified by signals applied to either or both upper or lower stage transistor bases. The overall current through the arrangement is set by the lower transistor, which is shared by the upper two; assuming both of the upper two transistor's bases are held at the same voltage, the currents will be shared equally; if one is raised with respect to the other though, its share of the current will rise, having stolen it from the other and vice versa (the total current stays the same). So, wobbling the lower transistor's base will change the overall current, an upper's base that in both upper transistors, complimentarily, the combined effect is multiplicative gain variation. (Conveniently, one of the signals [usually the audio] can be applied to and recovered from the pair of upper transistors differentially, although it is not unusual for them to be driven one-sided, the opposing base grounded.) The one remaining drawback is that the operating points of all the devices are moving around in accord with the control voltage applied to the base of the lower transistor and so the control voltage unavoidably appears as part and parcel of the derived output signal.



A. Cascade amplifier.



B. Long-tailed pair.



C. Gilbert cell or double-balanced multipler.

Figure 25-107. VCA design.

25.16.2.2 Gilbert Cell

Fig. 25-107C shows what is the essential heart of a good VCA—two long-tailed pairs back-to-back. Actually it's a bit more like three; a long-tailed pair with a long-tailed pair in each output leg. So universal is this basic configuration that it has become the Hoover of VCAs—it is what springs to mind when *VCA* is mentioned; variations and extensions to this theme are used extensively. Called variously the *Gilbert cell* or, by RF guys, a *double-balanced modulator* its main attributes are the innate cancellation in the output of both the applied signal and of the control voltage (CV); all that appears at the output is the product of the applied signal and CV. Product implies it is the result of multiplication, which it indeed is. The circuit is the basis of a good if hardly perfect analog multiplier. Better yet, since it uses good old transistors with their exponential base-voltage to collector-current response, the control law is for the large part linear with respect to decibels of gain and attenuation. Which is why all the bother and complication.

25.16.2.3 Log-Antilog Cell

A different approach, resulting in a different internal topology is what could be called the log-anti-log approach and is schematically described for simplicity (although the actual integrated implementation is nothing like this) in Fig. 25-108.

It again relies on the exponential relationship between the base voltage and collector current of the transistor. The first stage is a log convertor, converting the (positive-going in this example) input signal into a (negative-going) logarithmically representative voltage; summed in with this is the *control voltage*, which, since we're in log domain, is linear voltage per dB; the composite is then antilogged back into the real world. The effect of adding or subtracting the control voltage is to increase or decrease the linear end-to-end gain.

25.16.2.4 Commercial VCAs

Commercial IC VCAs typically use one or other of these approaches; VCAs are almost always acquired and used in IC form. One built out of discretes will work, of course, but the inherently much closer matching of active semiconductors on the same substrate reduces much matching of parts and tweaking out of various offsets, and the manufacturers have gone to the bother of thermally compensating and biasing everything up such that it "plays nice" with the real world. Nevertheless, for optimum operation of this arrangement or circuits based on it, preset adjustments are the norm, even for integrated versions. Stabilization of operation against temperature is a further complication, if perhaps less so for consoles that will spend their lives in air-conditioned environments.

Beyond the basic and remarkably well-performing circuit element, there are, of course, other issues that come along with real live electronics. A prime consideration is noise; the best operating points for the transistors vary depending on the parameter that needs to be optimal (see the discussion on microphone amplifiers for collector-current versus noise); what may be right for noise almost certainly isn't right for adequate large-signal handling. Attempts have been made to provide for both by altering the bias point of the transistors dynamically in accord with applied signals, such that they're closer to right for both the low-level region (noise) and high-level operation. Another approach has been to parallel up many of the IC VCAs in order to improve the combined devices' noise-voltage to noise-current ratios to improve noise and better suit operation at ordinary audio signal impedances and levels (this goes hand-in-hand with paralleling or using multiparallel input transistors to optimize OSI in transformerless microphone amplifiers).

Input buffering and conditioning of the control signals makes them easy to use; already quite linear, the linearity can be extended over a greater control range and can be arranged to be changed at so many dB per volt of control signal (say, 20 dB/V) as to be convenient with the A/D and D/A converters in an automation system and the voltage swing off dc-driven faders. Typically, though, the control port sensitivity on integrated VCAs can be much higher than this—a few mV per dB—and needs to be treated with respect.

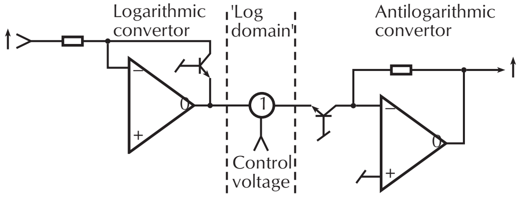


Figure 25-108. Log-Antilog VCA Principle.

25.16.2.5 Control Voltage Noise

This discussion highlights a crucial design issue—that of ensuring a very quiet control signal. This might seem an odd concern until one realizes that at typical control sensitivities, mere mV of undesired ripple or noise on that control line will modulate the through audio noticeably. Note *modulate*. Since the balanced modulator that is the VCA will not permit control voltage itself into the output, a real audio signal has to be passing through the VCA for this modulation to take place. This, more than anything, is the underlying cause of VCAs' largely undeserved reputation for sounding dirty. Like all these kinds of aspersions, there is a germ of a reality behind them. And this one is CV noise. Any circuitry involved in the CV should be handled with the care one would apply to the "real" signal path; there is a very real temptation to be more casual (and cheaper) with control stuff that should be avoided.

A less-than-obvious concern springs from the fact that the audio path and the control path are not only cross-disciplinary, but are architecturally dissimilar. Audio paths are (assuming mixer channels) following the signal flow, as are their grounds, while the control voltages for a number of, if not all, channels are being handled en masse and distributed star fashion. If ever there was an inadvertent recipe for a ground-induced noise problem, this is it. If the CV is referenced to a ground that is moving in any fashion at all in relation to the audio ground at the VCA, then that difference is effectively added to the CV as far as the VCA is concerned creating noise modulation.

25.16.2.6 VCA Channel Application

Fig. 25-109 shows a typical implementation of a high-end integrated VCA.

The THAT Corporation ("son of dBx") VCA-type 2180 is a current-in, current-out device for audio, hence, a standard current-to-voltage convertor using a good bipolar op-amp following. Note also a seeming overkill op-amp on the control-voltage summer. The control port is where things get interesting. The control feed to the VCA can be a summed combination of several different sources (a quick point here—since VCA control voltages are logarithmic, adding voltages results in multiplication in the VCA, or in other words, the dBs represented by the voltages add or subtract):

1. The channel fader, only it isn't, really. It's actually the output of a D/A convertor that is either reflecting the fader position as sensed by an A/D convertor, or replaying a prior fader position from the automation system. But, for now, we'll call it the fader.
2. Gain-reduction control from the channel dynamics. It is common to use the high-quality fader VCA as the gain-control element for on-channel dynamics. It presupposes the dynamics have deterministic feed-forward detectors and conditioners. Obviously, a feedback-style compressor could not use this VCA.
3. VCA subgroups. A common feature on sound-reinforcement consoles, these are controlled by a central set of a number of VCA group master faders (usually eight). These generate a control voltage,

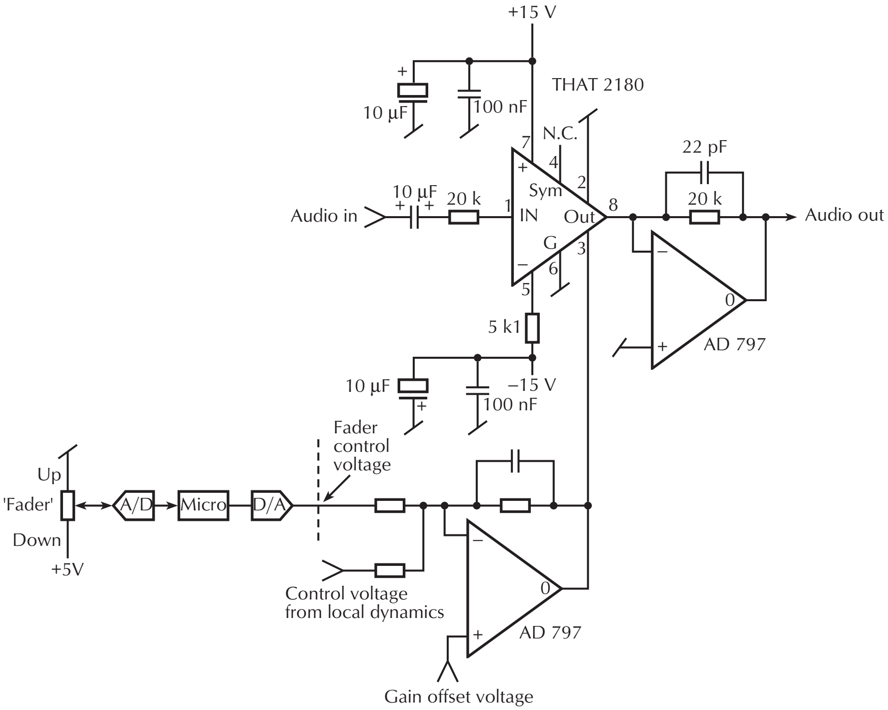


Figure 25-109. Simplified channel-style VCA using commercial IC.

each of which is bused up and down the length of the console; each channel has the option of selecting one (or more, if one likes danger in one's life) of these voltages to be summed into its own VCA. This is a very convenient manner of grouping related channels under a control without having to create real audio subgroups.

4. VCA master. Again, a centralized fader only operating as an overall master over all channels contributing to the main mix bus. Although seeming to be redundant being that there is almost certainly a real audio fader on the mix bus output, a VCA Master has the advantage that all the levels of sources contributing to a mix can be adjusted, rather than the output of the mix stage. Helps avoid headroom problems in the mix stage.

It should be stressed that the 5 V supply for the 0 to 5 V control signals should be oppressively regulated and fabulously quiet, squeaky clean. Borrowing some off the nearest micro and hanging 100 nF across it doesn't count—sorry.

In console designs with sophisticated computer control, all but the local channel dynamics control signal are manipulated and summed digitally and this composite result is fed to the channel VCA via a D/A convertor; this dramatically simplifies the multiplicity of summed analog control voltages per channel.

25.16.3 Digitally Controlled Amplifiers

VCAs are not the solution for all variable control in analog circuitry. In order to be driven from a digital control system a D/A convertor output needs to be used to derive an analog control voltage for each VCA. This can get very expensive, very quickly. A gain controllable stage that can be more directly connected to the controlling microcontroller is desirable.

25.16.3.1 Multiplying DACs

A more direct approach, meaning it can be driven directly off a digital control system, is to use a multiplying digital to analog convertor (MDAC) (don't you just love that "multiplying" bit?)—In particular, a referenced-input four quadrant multiplier, which implies it can produce an output both positive and negative in potential (or current in this particular case) and which is proportional to a voltage applied to its reference terminal, Fig. 25-110. Now, bear in mind these devices were never intended to be used in this way, but luckily they do so well.

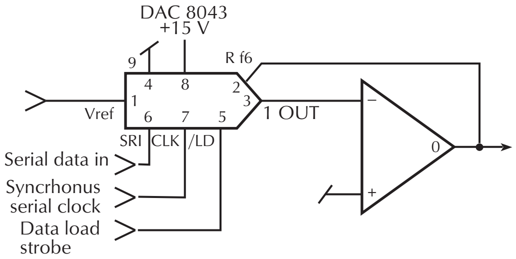


Figure 25-110. MDAC as rudimentary gain control

The audio signal is applied to the reference pin; a digital number, in this case 12 bits wide serially fed into the device, is applied to the 12 bit R-2R-style ladder DAC, Fig. 25-127; the audio signal is attenuated in proportion to the applied digital number with respect to the 12 bit maximum (1024 steps). The output current is sensed and converted back to a voltage by the following virtual-earth input amplifier, using the friendly internal feedback resistor around the op-amp. The interface is dead simple, linearity is pretty good, the signal handling is excellent, and the noise isn't bad—dancing in the streets!—except every time the gain is changed (a new digital word transferred into it) it makes a little tick noise, which is very audible on high-level signals, low-frequency signals, and especially the combination. In fact, as the gain is moved (a la fader), classic zipper noise is very evident. The only good news about all this is that for a large part, program material's spectral content masks this noise. However, when the device is used as a frequency or Q-determining element in an equalizer, the effect becomes comical; depending on one's sense of humor. There are two approaches necessary to nail this noise, since it is actually due to two separate causes.

25.16.3.2 Charge Injection

This is a near unavoidable effect in CMOS and other electronic switches, where a tiny amount of differentiated charge impinges itself onto the signal path from transitions of the control port. In a multiplying DAC, any number of bits may be changing as the gain is varied, and so the total charge injection varies correspondingly. It is, however, almost completely independent of the applied audio signal.

Cancellation works well, with reservations. One approach is simply to use a second MDAC with its own inverter that sums into the virtual-earth point of the main MDAC path, with its reference pin undriven. However, with only slightly more complexity the

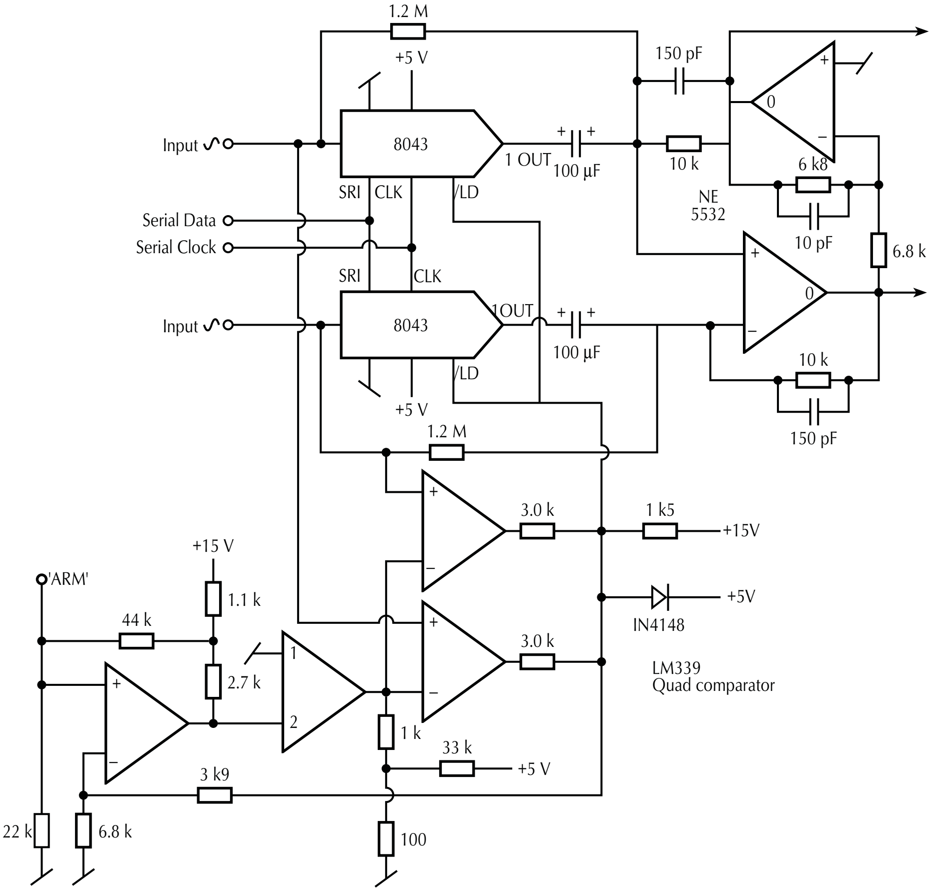

1.2 M
+5 V
150 pF
+
0
−
Input
8043
1 OUT
+
+
100 µF
10 k
6 k8
SRI
CLK
/LD
NE
5532
Serial Data
Serial Clock
10 pF
6.8 k
SRI
CLK
+
0
−
Input
8043
1OUT
+
+
100 µF
10 k
/LD
+5 V
150 pF
1.2 M
+
−
3.0 k
1 k5
+15V
+15 V
'ARM'
1.1 k
+
+5V
44 k
IN4148
1
3.0 k
LM339
Quad comparator
2.7 k
+
−
2
−
1 k
33 k
+5 V
3 k9
22 k
6.8 k
100

Figure 25-111. Differential MDAC gain control, with zero-crossing enable.

arrangement as in Fig. 25-111 emerges, which is our old friend the Superbal differential summing arrangement fed by both DACs being driven differentially. Not only does this provide a gain-control stage with full-rail differential signal-handling capability, but also the charge-injection noise is substantially canceled. To get the best noise cancellation, however, the DACs really need to be matched (a DCR test through the ladder is a reasonable guide for matching) or pairs on the same substrate employed.

25.16.3.3 Zero-Crossing

The second impulse noise cause is attempting to switch a high-value signal; any truncated or very rapidly level-shifted high-level signal is going to go "click!" (Run tone through a switch and turn it on and off a few times. The switch click—nicknamed *tone-click*—will vary in intensity seemingly at random; that's because the switching is occurring at random points through the sine wave's cycle. Those at or near the crests of the sine wave will click loudest.) The simple solution is, don't. If one arranges only to change gain while the applied signal is crossing through zero or is at a low level, this manifestation will all but disappear.

This is a control issue rather than an audio path issue; Fig. 25-111 illustrates differential MDACs with a peripheral circuit that achieves near zero-crossing. The MDACs in Figs. 25-110 and 25-111 are double-buffered. In other words, it is possible to load a new gain value into them without disturbing the current operational gain and then transfer the new value over when desired by means of the /LD control pin. The arrangement shown allows the controlling micro to do a "hit and run" on the circuit, depositing the new gain data and telling the circuit to take it at the next zero-crossing; the micro doesn't have to hang around waiting for a zero-cross to occur. It can be zooming around setting up other MDACs in the meantime or attending to other microlike things. The circuit is addressed, the new gain data serially clocked into the MDACs' first buffer and then the micro nudges high the ARM line. (The ARM line needs only an instantaneous +5 V pulse; positive feedback around the first comparator keeps it set. Likewise, it can if desired be nudged down to dis-ARM—nice, but useless.) Comparators wait for the applied audio signal to fall into a low near-zero signal window, at which point an instantaneous strobe pulse for the /LDs is generated, which latches the new gain data into the MDAC ladder and simultaneously cancels the ARMing.

This arrangement would not be the avenue best traveled for dynamics, being that it takes a comparatively long time to load in data and wait for zero-crossings, limiting apparent responsiveness—VCAs are a far better course for dynamics—but for anything else it works a treat. With reasonably matched 12 bit MDACs this gain-control circuit is virtually transparent and even works well in high-*Q* filters and EQs. It's still not inexpensive though, and the dawning realization of exactly how many of these circuits (or DAC/VCAs) would be necessary to fulfill complete automation of a decent size mixing console, and just how much they'd all cost, has quite a stunning effect.

25.16.4 Discrete Logic and Programmable Gate Arrays

In the next few (as in many earlier) pages, some interface circuitry will be described in seeming excruciating detail; the literal approaches taken will be valid for small or localized circumstances where discrete logic makes sense and the tremendous advantages (cost, board real estate) of integration into large-scale programmable digital parts cannot be realized; where the system is large enough, the detail serves as a road map for what needs to be emulated in the PLD (programmable logic device) or FPGA (field-programmable gate array). The ubiquity of these parts now has led to a hardware design approach that is at once bold yet somewhat alien to those who still remember tape-and-dot layouts; everything on a board, say switches, resolvers, converters, etc., are taken directly to pins on a gate array; the interface to the host microcontroller is brought to the gate array; then how it's all interconnected, strategized, timed, polled, strobed, etc. becomes a pure (software) programming exercise for the gate array. Errors and changes similarly become just software changes, too, not board re-spins.

25.16.5 Recall and Reset

Remembering the position of controls in a conventional console was the great innovative burst of the late 1970s. The niceties of techniques vary, of course, but Figs. 25-112 to 25-116 are reasonably representative. The great advantage of this sort of method is that it can be applied to an existing design with virtually no modification; all that's required is a rider pot on the back of variable controls (although this can be a bit difficult with dual-concentric pots) and an extra pair of contacts on switches.

25.16.6 Data Acquisition

The digital data-capture system is fairly straightforward. Switch closures are sensed in batches of 8 (or 16 if a large microcomputer or a minicomputer is in use), while each individual pot position is resolved to the accuracy afforded by an 8 bit analog-to-digital (A/D) converter—256 possible positions. Although very high for resolution and practical resetability of most pots, it is actually harder work reducing the capability than leaving it be! This may be true for pots, but with high-quality faders it may seem too coarse; 12 bit resolution may be necessary.

Two different types of input multiplexers are needed, one for switch closure sensing and an analog switcher for the rider-pot voltages. In computer thinking, each set of eight switches and each rider-pot is regarded as a single memory address; an entire console worth of control settings occupies a chunk of the computer memory map. It's easy for the processor to run through these addresses and collect a set of data.

In Fig. 25-112 a channel's worth of multiplexing is shown—32 switch sensings and 16 pots. Inexpensive CMOS switchers are used throughout; speed isn't a real problem. The switch-sense multiplexers directly hit an 8 bit data bus, which can either be the actual processor data bus (if the processor clock speed isn't too fast for the CMOS propagation delays) or, ordinarily, a buffered sub-bus with slower timing. Speed freaks wondering why things are almost deliberately slowed down should remember two things:

1. A data acquisition system such as this running at even a leisurely processor clock rate is quick! This is not a real-time variable system, it's intended mostly for snapshot storage of console status; the acquired data is really trivial by most processor system standards.
2. The A/D conversion time for the pots keeps the processor hanging about in wait states far longer

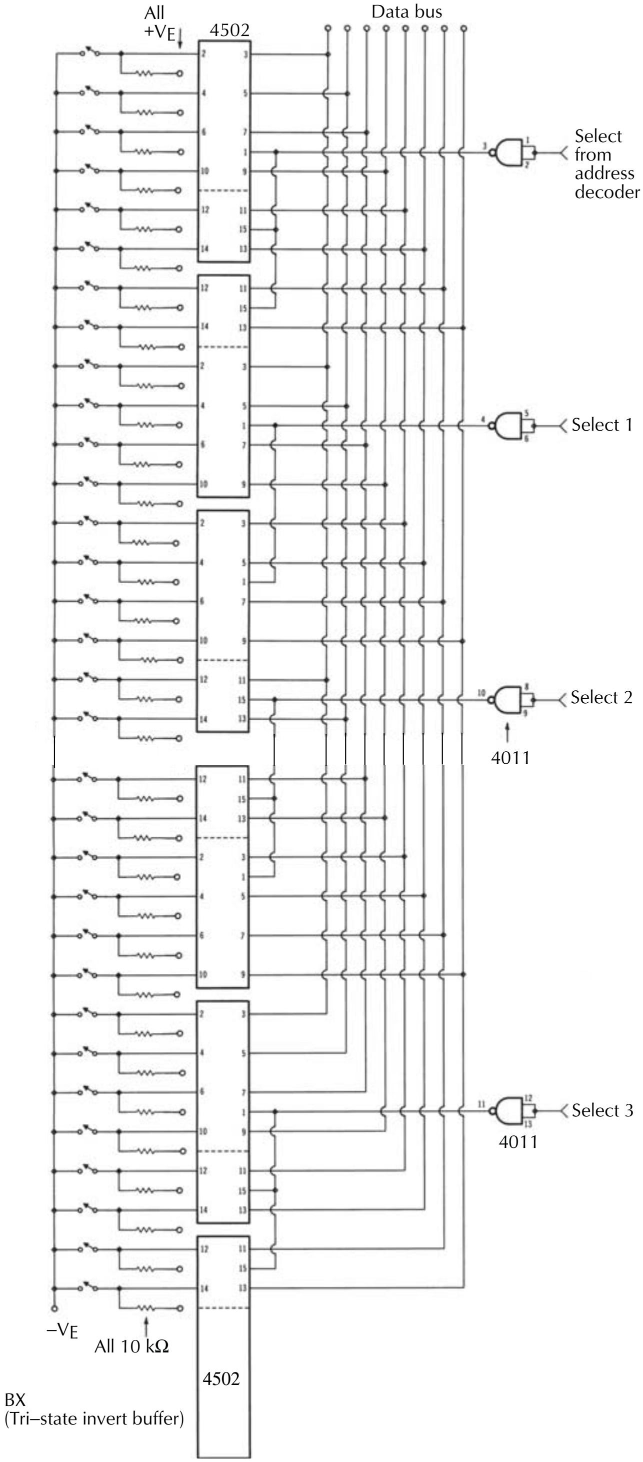

All
+VE
4502
Data bus
Select
from
address
decoder
Select 1
Select 2
4011
Select 3
4011
−VE
All 10 kΩ
4502
BX
(Tri–state invert buffer)

Figure 25-112. Rider switch multiplexing.

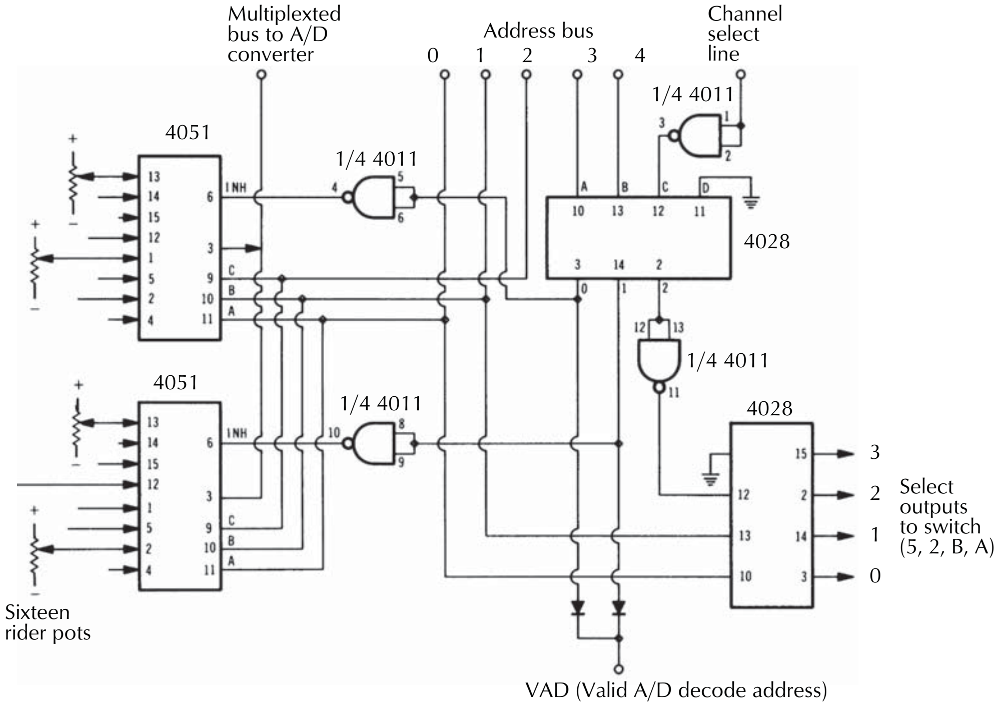

Multiplexted bus to A/D converter
Address bus
0 1 2 3 4
Channel select line
1/4 4011
4051
1/4 4011
INH
4028
1/4 4011
4051
1/4 4011
4028
3
2
1
0
Select outputs to switch (5, 2, B, A)
Sixteen rider pots
VAD (Valid A/D decode address)

Figure 25-113. Rider pot multiplexing and address decoding.

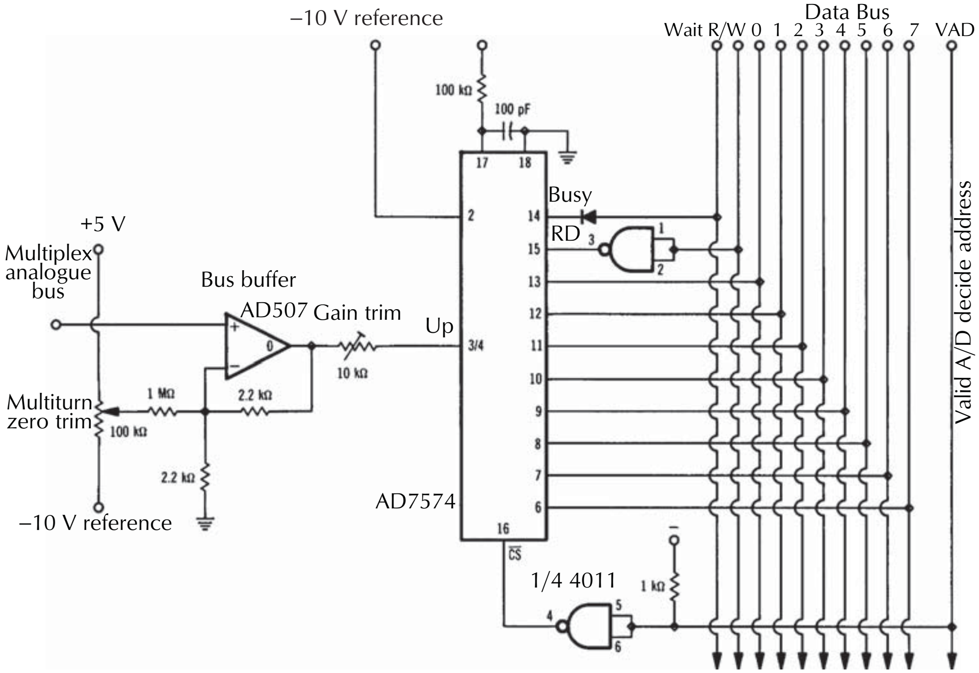

−10 V reference
Data Bus
Wait R/W 0 1 2 3 4 5 6 7 VAD
100 kΩ
100 pF
Busy
RD
+5 V
Multiplex analogue bus
Bus buffer
AD507
Gain trim
Up
10 kΩ
Multiturn zero trim
1 MΩ
2.2 kΩ
100 kΩ
2.2 kΩ
−10 V reference
AD7574
CS
1/4 4011
1 kΩ
Valid A/D decide address

Figure 25-114. An A/D converter (as part of the system in Fig. 25-113).

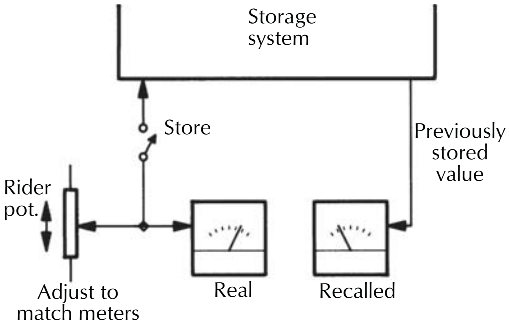


Figure 25-115. Simplistic manual reset system.

than a single address acquisition by a typical imbedded microcontroller.

A/D conversion can be done in a number of fashions for this system.

25.16.7 Control A/D Conversion: Central or Distributed?

Central conversion means that there is just one A/D converter in the processor rack frame. All the multiplexed rider-pot voltages hit one bus, which is then A/D converted centrally and the result of the conversion goes directly to the processor data bus. The obvious advantage is low cost—only one converter. Disadvantages are speed (a successive-approximation converter takes several processor clock cycles to perform a conversion; this can be obviated by using a very high-speed comparator-type flash converter) and bus slewing (caused mostly by bus capacitance). Since each rider-pot source is not zero impedance unless it's at one end of its track or the other and CMOS analog transmission gates have a finite on impedance, there is a definite time constant involved, with the bus capacitance needing a certain amount of time to charge to the correct potential (as determined by the rider pot). The previous bus potential can, of course, be anywhere depending on the previously selected position of the pot. Even if this time constant can be made short with respect to an acquisition cycle, it even then really can make a nonsense of 256 level, 8 bit resolution!

Buffering each rider pot to present a known zero impedance to the multiplexer is a partial solution; buffering the multiplexer output—a seemingly obvious solution—creates more problems. First, the buffer output needs to be gated away from the bus for the times it's not addressed, so there is a transmission gate impedance there regardless. Second, it has to be a very fast follower if it isn't to create worse slewing than the bus! Remember that the multiplexers are switching at processor or sub-bus speed. Suitable amplifiers tend to be as expensive as they are fast.

Distributed A/D essentially means having a converter on each subassembly (channel) or for a small number of pots. There are, for example, proprietary converters that continually read and cycle through eight inputs independently of the main processor, yet allow free access to the collected data. Either a batch of these or the equivalent built from individual converters and multiplexers allows the processor to work unhampered by conversion-related hangups, while also keeping all system interconnections digital. Similarly, this method avoids gross bus-slewing (there would no longer be a long analog bus). As always, there are difficulties:

1. Although 8 bit successive approximation A/D chips are now cheap, the number has grown.
2. There are more bits on the channel subassembly.
3. The multiplexers feeding the on-board converter are still switching at high speed—slewing inaccuracies are still possible. Clever priming algorithms can increase conversion accuracies while maintaining a high overall acquisition rate, almost as high as for switch-closure bytes. These set in motion a conversion on one channel, allowing plenty of time for switcher settling and so on before the result is looked for on the bus. During the idle period the computer is dealing with other setups and results from other channels.

Both central and distributed systems are successfully used in circumstances where ultimate speed isn't that important. Remember that this is not intended as a real-time application and the actual amount of data is small. Accurate enough resolution is reasonably easily achieved.

With the cost of such devices becoming pocket change, it is not unrealistic to throw a microcontroller at each channel simply to perform these tasks; a significant reduction in parts can be afforded.

25.16.8 Recall Display

Figs. 25-114 through 25-116 describe how all the relevant console control positions can be digitized into processor-manageable form. Storage on a mass medium such as hard disk or network is a fairly simple computer file-management exercise, as is recalling it. What to do with the recalled information is now the question.

It is assumed that this particular requirement is informational recall only, not hardware reset (i.e., setting up the parameters of the channel to their stored values). Eyeball comparison and human tweaking is the resetting mechanism employed. The comparison is between a recalled value displayed on a meter, LED column, bar display, null indicator, or up on a GUI (Graphical User Interface) screen, and the immediate real value read from the control in question and displayed on an adjacent display. As the relevant control is tweaked, its indicated value will be higher or lower than the stored value; when the two are matched, then the control position is the same as it was when the snapshot was taken. Fig. 25-115 shows in simplistic form the basis of the matching process.

GUI displays are presently the easiest way of performing this matching. So much information is visible at once, which is a blessing in this circumstance.

Even with increasingly common totally programmable/recallable consoles, screen-based display of control statuses is a very useful function, if in addition to localized feedback to each individual control. It looks good, too.

25.16.9 Nulling

Null indicators are particularly easy to implement and use. They usually take the form of a pair of LEDs adjacent to the relevant control. If the real value is higher than the recalled value, the upper LED lights; if it is less, the lower one lights. If they both come on, the two values are matched. Even simpler nulling indicators take the form of a single LED that only comes on (or alternatively goes out) when the two values match. A nicer arrangement is a single-cell green/red LED giving an unequivocable "go" or "no go" indication. This device makes it particularly easy to spot anything out of order on a channel.

A fairly elaborate demultiplexing system has to be plumbed onto the channel board, however, to deliver the software-derived nulling indications to the front-panel LEDs; Fig. 25-116 is representative. A further amount of processor memory area needs to be dedicated for this output facility in addition to that already spoken for by the input multiplexing.

Another software consideration when using null indicators is that the chance of actually finding the 1 in 256 position that is correct is pretty slim. Reducing the effective resolutional accuracy to fewer bits can make the operation a lot simpler. The accuracy of the recall system can easily outstrip that which it is monitoring and it is a judgment call between resetting precision and the ease of so doing.

As laborious as these facilities may be operationally, a complete reset of console parameters can be achieved. It is considerably less laborious and inaccurate than writing everything down.

Interestingly enough, any bus-slew inaccuracies engendered on storage tend to be canceled during recall. When all the controls on a channel are reset at or close to their original settings, all the bus errors will be very similar to those present when stored.

25.16.10 Resetting Functions

The next logical step in developing computer assistance is for the machine not only to remember console settings but also to reestablish the console to its previous operational state on command. This means that if the multitrack routing on channel 27 was going to machine track 15 when the console status was stored, then regardless of what has happened or how the routing may have altered or configurations changed, upon recall channel 27 will go to track 15.

Most of the circuitry described in this chapter, especially the multitrack console example, is from a generation of design where active resetting of all major switched functions was a requirement. Variable controls were not even considered as candidates for resettability since it demanded too great a shift in technology. As we will see, the techniques necessary for that become instrumental in a deeper, broader change of console design, structure, operation, and philosophy.

Every switched function that is intended to reset needs to be made electronically controllable; the techniques are detailed in earlier sections. This replacement, by and large, has already been implemented with other ends in mind, such as simplifying PC layouts, avoiding large physical switches, and, not the least, facilitating some of the tortuous signal rerouting required in a modern production console.

In addition to the data acquisition system as just described (i.e., switches to computer), a second digital distribution system—computer to switches—is required. Techniques similar to those described for nulling indication work well.

25.16.11 Motorized Pots and Faders

Motorized pots and faders look and feel like conventional ones, only a clutched motor drive controlled by a servo allows the mechanics to be reset to any point on their travel. A rider track, either a normal resistive track encoded by an A/D converter or a digital track direct input allows a microcontroller to keep track of the posi-

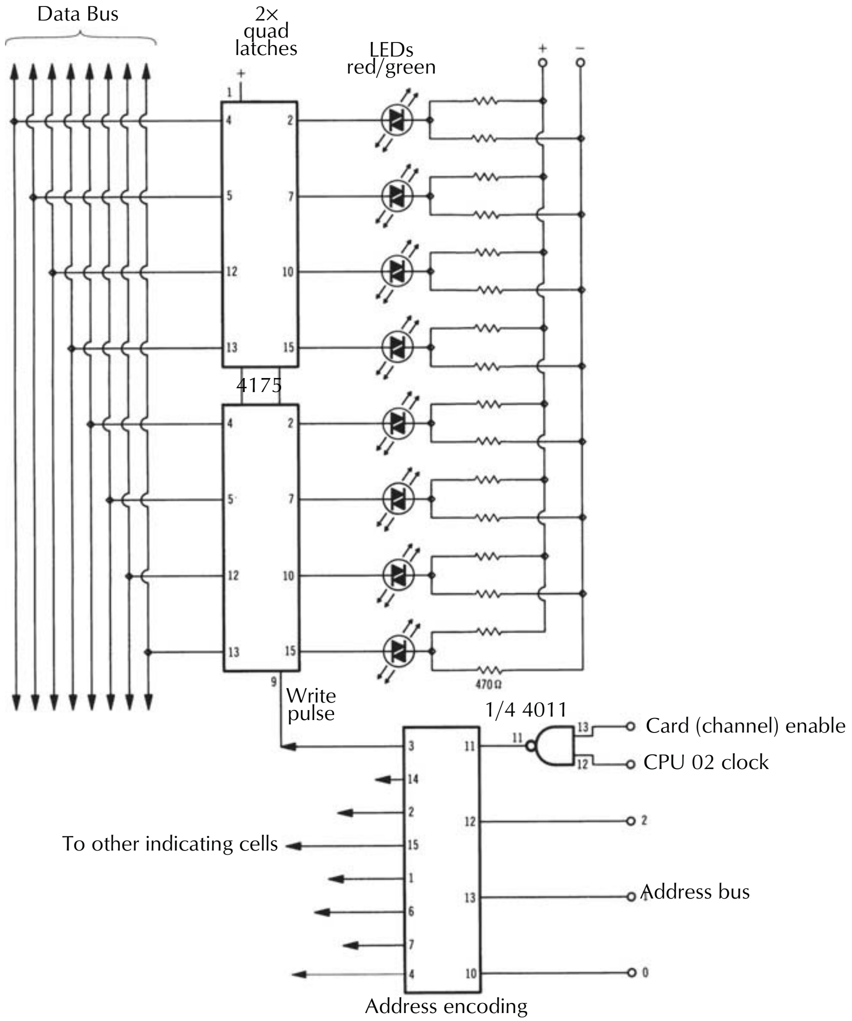

Data Bus
2×
quad
latches
LEDs
red/green
+
−
4175
470Ω
Write
pulse
1/4 4011
Card (channel) enable
CPU 02 clock
To other indicating cells
2
Address bus
0
Address encoding

Figure 25-116. Nulling indicator decoding.

tion. Comparing its present position with one previously stored drives the servo to equalize the two—i.e., return the control to its prior position.

These are increasingly used in newer consoles, particularly automated or soft consoles, where one physical control can be responsible for many channels or functions.

25.16.12 Resolvers

Resolvers are continuously rotating (no end stop) controls that otherwise look like a conventional potentiometer. Indication for these is commonly arranged to be a circularly disposed set of LEDs around or within the resolver knob rather than linear, adjacent. Such arrangements of varying degrees of cleverness are a staple of control surfaces nowadays. A resolver, when rotated, sends out two streams of pulses, half overlapping as in Fig. 25-117; in other words, they are 90° out-of-phase or in quadrature. This is enough information to determine not only how fast it is rotating (by counting the number of pulses from one of the trains) but also in which direction. These two, rate and direction sense, are enough for a controlling processor to analyze and appropriately perform control.

The simple circuit of Fig. 25-118 sorts it out; it's a 4013 D-type latch. The data port is fed by one train, while the edge-triggered clock input is fed by the other.

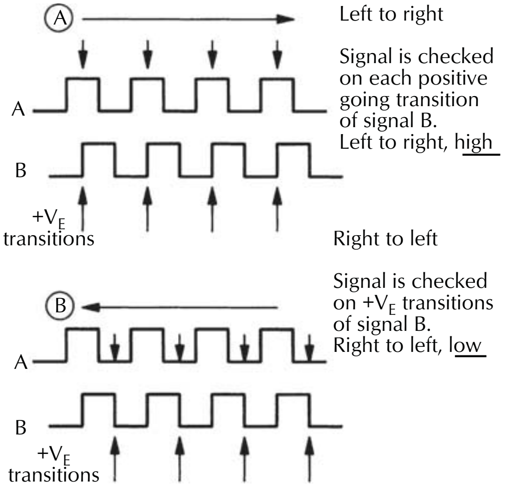

Ⓐ
Left to right
Signal is checked on each positive going transition of signal B.
Left to right, high
A
B
+V_E transitions
Right to left
Ⓑ
Signal is checked on +V_E transitions of signal B.
Right to left, low
A
B
+V_E transitions

Figure 25-117. Quadrature streams from resolver indicating how speed and directions are sensed.

If the clock is triggered by the rising edge of the A train and the B train is active, then the latch output goes high, indicating one direction of rotation (left to right in Fig. 25-117). In the other direction, the rising clock edge from A corresponds to B being inactive, so the latch output goes low.

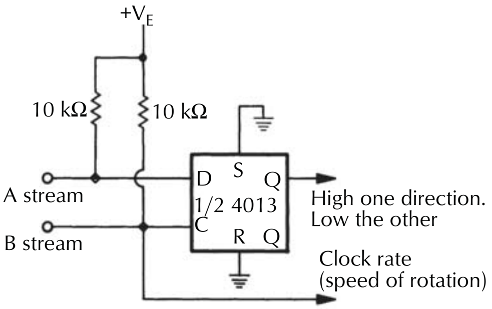


Figure 25-118. Resolver decoder (using a D-type flip- flop).

It is rather a simplistic circuit that assumes that the making contacts of the resolver are perfect and no false triggering will occur. With more swanky optical resolvers this may be true, but with mechanical ones a little debounce clean up prior to the D-latch gates may be advisable.

25.16.13 Control Surfaces

A large problem with recording and live consoles has been precisely that—they've gotten large. Console channels have grown into long, thin strips for purely historical reasons, and the manufacturing technique of hanging all the signal-path electronics on acres of dense PC card has just tagged along with little evolution. Removing the audio electronics (analog or digital) from the control surface into a remotely controlled equipment rack seemed quite an obvious development, although until recently it was a technically unwieldy one. Many types of analog circuits lend themselves to direct remoting. For example, VCAs for level control need, in essence, a single dc control line. Others, such as equalizers and microphone preamplifiers, don't. Noise and difficulties in extending nonzero-impedance configurations are both significant problems. As with everything else, these areas of difficulty look quite different given a dose of digits. As has been seen, with only minor compromises, digitally controlled remotable audio circuits of all sorts are realizable at some cost and complexity; it is entirely possible for the control surface to become now just that, no audio need go anywhere near it.

The question of whether the control surface and signal processing electronics should be divorced and live in environments possibly better suited to them is a more vital one with digital mixers than with analog; although possible, it was (is) actually very expensive to do remote, fully digitally controlled analog (DCA) circuitry that sounded decent and wasn't riddled with clicks, burps, and fizzies. It is more expensive now, amusingly, than having a fully digital signal path, which has really sorted that argument out once and for all.

25.16.13.1 The Single Channel Concept

There is an immediately apparent redundancy with large consoles—rows and rows of identical channel modules. The first intuitive step would be to reduce all those to just one set of channel controls that is selectable or assignable to any channel that needs tweaking. The first modification to this rather simplistic single-channel console concept is that the main level faders need to be kept continuously available in front of the operator; a button adjacent to each of the individual faders (the "ME!" button) calls the set of assignable channel controls to the channel to which that fader is related.

The second modification concerns the assigned controls. Like the knobless fader, they have to be separately acting for indicating. On being called, the indicating part of the control adopts the settings pertinent to that channel; the control, whether it be knob, switch, or fader style, can then act on the selected channel with the indicators following their action on the remote circuitry.

A wide and glittering variety of controls have evolved to suit this requirement, but they all basically have a row or concentric ring of indicating lights, fluorescent indicators, or LCD panels disposed around the digital resolver control knob. Alternatives to knobs, switches, and indicators such as interactive GUI screens suffer from an ergonomic disconnect between the physical operation of the control and from where the relevant feedback is displayed, unless the control becomes a mouse-driven widget on the screen. GUIs do, however, pose a very attractive supplemental display method, if not primary.

A third modification to the initial rationalization concerns the many auxiliary mixes found in a console, whether they be for effect feeds, foldback, or, perhaps most importantly, multitrack monitoring. Although the controls for these are traditionally regarded as channel controls, intuitively they are thought of and operated on horizontally across the console; if someone's setting up a foldback mix they'd most likely be working along the row of controls for that mix bus (to which they'd also almost certainly be listening via monitoring) and have very little interest in any other channel controls at all. Making the operator select each channel at a time to do such a routine mix setup is a very retrograde move—it imposes an unwelcome multi step process that diverts concentration from the task at hand to the means of achieving it. Quite sensibly then, any same-function bus-oriented controls should become accessible together. This is precisely the rationale behind the channel faders all being accessible simultaneously. Ideally, a row of interactive knobs, one per channel across the console, the function of which follows the feature of interest (like that foldback mix) is appealing. Such have been variously called smart bus or virtual controls. A neat bit of further rationalization comes into play here; the consolewide set of controls implied in having parallel access to auxiliary buses (meaning that in addition to a fader for each channel there would be an auxiliary bus control also) can be avoided if really necessary by using the already existent faders. After all, if we're busy setting up an auxiliary mix, we won't be overly concerned about other mixes, including the main one. Even if something does need instant attention, reassigning the faders to main mix is only a button away.

So here is the essence of control surface rationalization. There would be a row of moving faders and possibly a smart knob, one for each channel, with an adjacent control select button (ME!) that renders a singular set of channel controls (whether glass or physical) operative on that particular channel. We would also have a row of buttons (with again possibly GUI supplementation) that selects on which mix-bus(es) the fader row is acting.

Early practical experience with this showed, even with operators who came to grips with the single-channel concept readily, that there should be ideally more than one set of channel controls—it is a common requirement to play two or more channels against each other in a mix. Secondarily it was felt that having a set of controls always set up on the one critical channel in a mix (the money mic) and having one or more floating surfaces perhaps represented a better compromise. This represents the fourth major modification to the single-channel concept, although most rationalized designs still lean to just supplying the one set of channel controls.

The great beauty of making all controls transient (i.e., not totally dedicated to any one channel's function) is that all console functions are implicitly digitally stored, recallable, manipulatable, and automatable.

25.16.13.2 Commercial Console Control Surfaces

There are, however, very strong reasons for retaining the single control per function and module strip layout familiar from big analog consoles of yore over to digital consoles, regardless of the undoubted temptation to rationalize.

By way of prime example, Figs. 25-119 and 25-120 show a world-class analog production console, the SSL Duality, and the corresponding world-class digital production console, the SSL C200—if you can tell them apart—which is the whole point. The very sensible argument is that there is a very large user base familiar with the ergonomics and accomplished in the use of the analog console, so why on earth force a learning curve on them? Both of these consoles are intended to work in a highly integrated fashion with/as digital audio workstations, explored below.

At the other end of the rationalization spectrum is the Innova-Son Compact, shown in Fig. 25-121. This is as close to the single-channel concept as it is possible to get; other than the faders all controls are centralized, following the ME! button's activation on the desired channel. There are some very clever features not obvious from the photograph: all the faders are moving faders; if a channel ME! button is pressed, all the group faders move to represent that channel's feeds to each of those groups. If a group ME! is pressed, all the channel faders move to represent each of the channels' contributions to that group. A very impressive surface and fun to drive.

It is not a surprising leap to see that a rationalized surface such as that can have the input (and output for

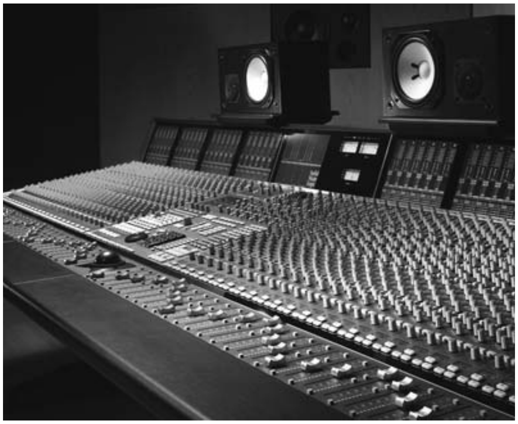


Figure 25-119. SSL Duality analog Mixing Console. Courtesy Solid State Logic.

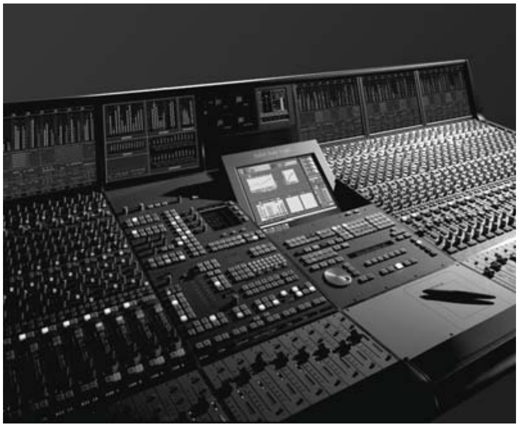


Figure 25-120. SSL C200 Digital Mixing Console. Courtesy Solid State Logic.

that matter) channels paged; this means that a switch can instantly throw a whole second (or more) batch of controlled channels up onto the surface. Superficially a great idea, since a modest-sized surface can drive a much larger console, this seemingly facile addition is far harder to come to grips with operationally than rationalizing the channels themselves was. It is quite unnerving to have half a console disappear! It takes considerable effort to design and engineer a surface with enough clues as to the background channels' existence and well-being to make paging a comfortable operation.

Fig. 25-122 shows a highly considered control surface design somewhere between the two extremes of knob-per-function and completely rationalized. The Wheatstone D-12 television audio console has centralized EQ, dynamics, routing, surround panning, auxiliary and mix-minus feeds, which are brought into play by,

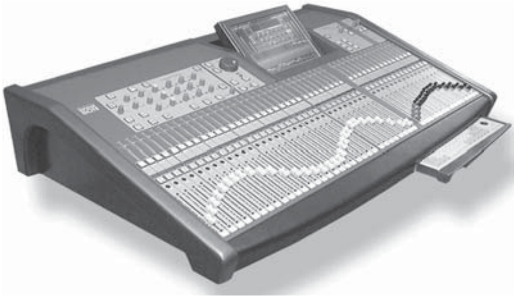


Figure 25-121. Innova-Son Compact sound-reinforcement console. Courtesy Sennheiser USA.

guess what, a ME! button on each relevant channel. Additionally, though, it is to be noted that there remains a considerable amount of localized control on each channel; these controls are what an operator needs to get his or her hands on immediately (and which of course can differ between setup and when on-air contexts), with no intervening selection step involved. (Remember, broadcast is a high-stakes no-second-take environment). Input metering is adjacent to each fader, and two sets of channel ID indication above each fader help assuage paging concerns; full console status and metering are spread across the numerous GUI displays in the penthouse.

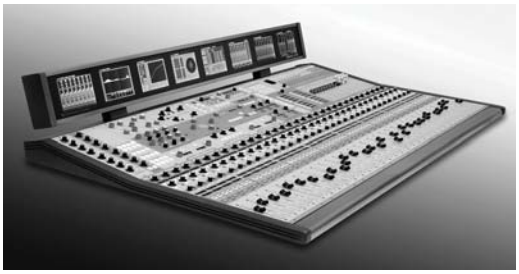


Figure 25-122. Wheatstone D-12 live console. Courtesy Wheatstone Corporation.

There are circumstances where the use of a room might be quite diverse over the course of a workday or likewise the technical adeptness of the users; in mind is that of a radio broadcast studio. One could have the problem of there being a perceived baffling sea of knobs for a disk jockey, yet insufficient control for a commercial producer. A convenient solution, falling out of the soft control surface concept, is shown in Figure 25-123; the hardware surface is very basic —just what an on-air presenter needs—but the (removable if need be) screen can be ME'd and mouse driven to have a full set of EQ, dynamics, and effects per channel: happy advertising producer. It would even be possible to run the console entirely from the screen, with no regard or need for the hardware surface.

Figure 25-123. Wheatstone Evolution 6 Console. Courtesy Wheatstone Corp.

And so in a few short paragraphs, we've moved from a knob per function to no knobs at all. Already in the fairly young game of control surface design sans frontiers manufacturers are rightly taking the measure of their cliental and producing surfaces much closer to their actual requirements than ever before, liberated by digital control. It is very encouraging. There is no universal perfect control surface solution; seemingly polar protagonists of the knob-per-function and fader-and-a-button approaches are equally exactly correct.

25.16.13.3 Control Surface Intelligence

Even if there is no signal processing going on in the same box, the control surface still has an awful lot going on inside it, Fig. 25-124. Typically there will be a large embedded controller, or even a PC-style microcomputer to administer things such as control surface host; it will likely be of the x86 persuasion, or a capable

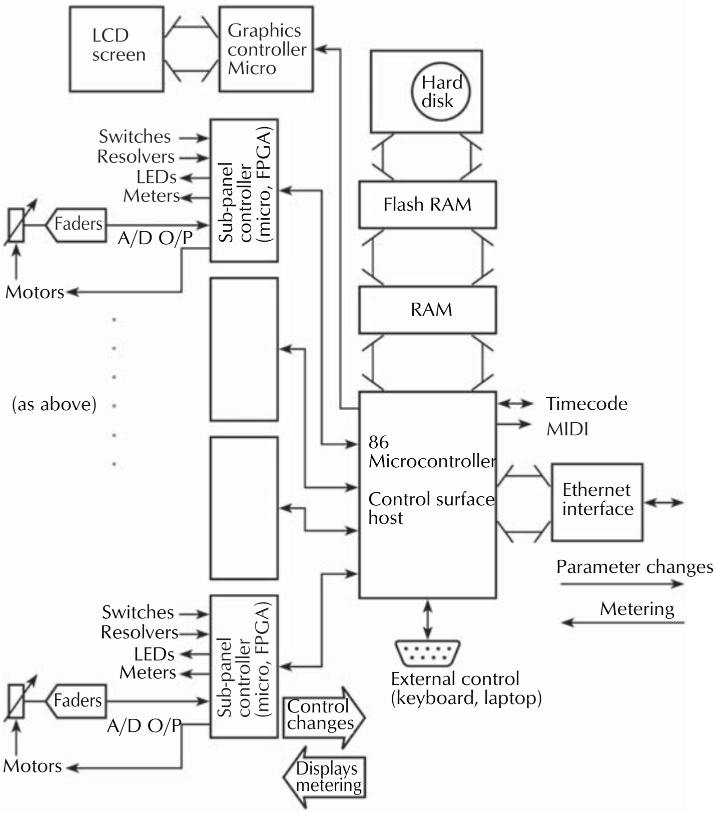

LCD screen
Graphics controller Micro
Hard disk
Switches
Resolvers
LEDs
Meters
Faders
A/D O/P
Sub-panel controller (micro, FPGA)
Motors
Flash RAM
RAM
(as above)
Timecode
MIDI
86 Microcontroller
Control surface host
Ethernet interface
Parameter changes
Metering
External control (keyboard, laptop)
Switches
Resolvers
LEDs
Meters
Faders
A/D O/P
Sub-panel controller (micro, FPGA)
Motors
Control changes
Displays metering

Figure 25-124. Control surface control architecture

embedded processer such as an ARM. Being short of the tyranny of thin module strips allows the garnering of a reasonable number of controls' data and the driving of a reasonable number of indicators without the indirection and bottle-necking encumbrance of a console long data busing scheme. Even a large console's worth of controls and indicators is easy pickings for a large processor treating them as medium-speed peripherals through industry standard buffers or FPGAs. Should a devolved scheme be necessary from a semimodular or macromodular approach to the surface each submodule may be looked after by a smaller embedded micro or even an FPGA, with communication from each back to the host by fast serial link.

Chances are, this host will also be feeding data to a subsidiary LCD screen driver processor (or two, or three), talking down Ethernet to the signal-processing host sending it fresh parameters (or even coefficient sets if they're being calculated at the surface end), receiving back from the processing host packets of metering information to be divvied out to the appropriate displays, and last but certainly not least attending to the level of automation (static snapshots or real time) in which the console is operating. To this end, it almost certainly has mass storage, such as loads of flash *ram*, and/or a hard drive.

25.16.13.4 Multiuser, Multifunction

User arguments can run something like, "But we might want to change several things at once, and Fred the producer likes to look after the monitor mix while I do the rest." The control software would naturally allow simultaneous control actions on a pair or across a group of channels to be ganged, which is fairly trivial and not the point being addressed. The main engineer console can be regarded and would be regarded by the host computer in console systems as simply a terminal, albeit the main one. There is nothing to stop other terminals of greater, equal but probably lesser or deliberately limited facilities having access to the main body of electronics, sharing the network and its resources, in other words. In practice they would have access to and be able to manipulate a preprogrammed subset of the total capability (e.g., our producer friend's monitor mix) concurrent to the main terminal or control surface. Another obvious secondary terminal would be a second or even third set of assignable channel controls for multi-op situations, although we can't help wondering how often they would be redundant except in the all-hands on-deck film mix-down world. As a capability it would go a long way to soothing the frustration of engineers new to the concept who are wary of losing so many controls at once in sacrifice to the new false god rationalization!

Simultaneous access to the same set of information is what the term *multiusers* is all about. Multiple control surfaces pose no real issues—control systems and networking operate so quickly in relation to the rate of changes a mixing engineer can make that several concurrent operators sense no interaction at all.

In computer terms the system described bears more than a passing resemblance to a hardware-related database, remotely controlled by a terminal or terminals down a network. Again, in computer terms, it's a pretty small database, and at least on the control side a pretty lightly loaded network, too.

25.16.14 Goodbye Jackfields, Hello Routing

Considering that one of the easiest audio subsystems to organize using digital technology is signal switching, it's astonishing jackfields still exist. Analog switching matrices are now at such a level of development that they can be considered transparent to the system. Digital routers of course have no impact on signal quality whatsoever. Neither in any way, even when many are cascaded, create performance limitations. They are dense (many thousand source/destination crosspoints will fit in the same rack space as 144 jack holes) and decreasingly expensive—much less expensive per crosspoint than a comparable jack circuit. Control is soft and the operation can thus be anything from a humble computer terminal, PC application, to effectively complete seamless integration as part of the console's control surface. Of course, within assignable systems, the matrix is controlled by an interactive control surface by the operator, all routings and parameters being storable, recallable, and resettable as are the rest of the parameters of the console, in real time if desired. Try that with 50 patch cords!

Inputs and outputs of everything internal to the console (equalizers, dynamics sections, front-end amplifiers, line-output amplifiers, and so on) and everything external to the console (effects, machine input and outputs, and so on) all appear as sources or destinations on the matrix. The concept of insert point has disappeared; anything can go in anywhere. After decades of things getting more complex suddenly things have become simple again—there is no system and no prewired interconnections. A system to fit a given circumstance is built up from scratch using all the circuitry building blocks interconnected as required via the matrix. A repertoire of usual starting points—preassembled patches—is stored and recalled as needed.

In pure digital signal-processing systems, this is taken a step further, where processing elements can be arranged in order at will or dropped into place in the form of plug-ins. No arbitrary system: full circle.

25.16.15 Integrated Control of Outboard Gear

A great many bits and pieces of outboard signal-processing gear (known vernacularly as toys) are involved in the successful production of present-day program material. Already the term o*utboard* is flimsy since via the system matrix, or via *plug-ins*, their signal paths are already firmly internalized. The old music-industry serial communications link MIDI, despite its limitations, still bears integration into any studio interactive system. The centralized control point for these is the interactive main control surface for the operator, and a MIDI-controller application is required unless the console as a whole (strangely but sometimes quite sensibly) merely becomes a MIDI slave to an external controller.

It is no coincidence that major players in the DAW world, and hence with influence tendrillike into the rest of pro-audio, were initially strongly into the instrument, machine-control, musical synthesis, and arrangement world of MIDI control (e.g., Steinberg, with CuBase); it helps explain, too, why there is such a tight integration of MIDI music-making capability with audio processing in these DAWs, and their look and feel is unmistakably MI (musical instrument) in flavor.

At last the impossible studio system of a mere few years ago, integrated, completely automated, and resettable in real time in conjunction with effects, storage machinery, and other systems such as video, is here.

25.16.16 Digital Audio Workstations (DAWs)

Control by GUI only, where all audio functions are controlled by mouse activation of on-screen widgets in the form of pseudo-knobs, buttons, and sliders, has been a natural progression, if for no other reason than it is cheap—there's no physical surface to build or buy! A GUI, though, presupposes that the actual signal processing is already in digitally controlled form, usually pure digital. Although a GUI can be part of an embedded system controlling a traditional digital console as described later, often it is part and parcel, along with the control code and signal processing code, wrapped up within a PC. This does not make it a nonconsole, all the parts and processes that make up a console are in place, just in the one place.

DAWs rapidly transcended being the dinky two-track editing tools they started out as and have become the de facto console experience in many spheres of audio. Characterized at heart by being (or having at least the appearance of being) software applications that run on the familiar PC or Mac; by absorbing the recording into the PC's hard drive, by providing access to just enough audio signal processing, by rationalizing control extensively so that it fits adequately on a screen, DAWs rule the nonlive and production audio arenas. While many DAWs totally run within and skirt the processing limitations of the host PC (which are becoming less limited as PCs become more capable daily) in some cases extensive additional DSP farms are employed to do the heavy lifting, leaving the PC mostly to do the user interface. In either case, the PC-based DAW is a perfectly valid multitrack production environment. Paradoxically, that which was the DAW's initial strength—the convenience, familiarity, and low cost of the PC environment and GUI—is now the major (there are others) drawback. Screen-based DAWs using point-and-click are highly rationalized in operation, and do not lend themselves well to other than single-operation-at-a-time usage.

This is the predominant reason DAWs and their underlying technology are still eschewed in any live audio activity, with more traditional (including rationalized!) console surfaces maintaining favor. That said, there are burgeoning after-market and own-brand control surfaces expressly to augment and improve operation of DAWs, and many traditional console manufacturers have embraced the underlying technologies and merged the two approaches quite seamlessly; these range from a small surface of little faders all the way up to major surfaces such as the SSLs above.

25.17 Digital Consoles

It is an impossibility given the nature of this book and the space available to give a thorough treatise on digital mixers and their techniques. It gets pretty mathematical, pretty scary, pretty quickly. What is intended here is an outline of typical audio digital signal-processing considerations, methods, and limitations from an intuitive and practical standpoint and ultimately in the context of a practical digital console design.

An analog versus digital divide still exists simply because as with any pair of such disparate technologies, what is easy in one can be hard in the other and vice versa; digital can do some things that are practically impossible for analog—time-related machinations, for example, which are typically gruesome in analog. Tritely, it used to be said that real-deal EQ and dynamics were the province of analog, being that it has hitherto been easier and cheaper to achieve nice-sounding, complex, and flexible phase and frequency response shaping with a handful of analog components; this has become less of a black-and-white proposition though as the size, speed, and power of digital signal processors have increased and relevant expertise and ears were applied. Very fine digital EQ, dynamics, and effects are indeed possible. Suggestion otherwise "is fightin' words" and many would suggest that digital audio processing has now surpassed analog in all important respects.

Particularly in mixing, switching, and routing there has been a dramatic bipolar switch over to digital purely on ease and inexpensiveness of implementation as appropriate parts became readily available; Fig. 25-124, a photograph of a couple of LSI ICs and a handful of support parts, illuminates this blindingly; of course it could be argued that the same could simply be achieved in analog with a mere 144 op-amps and 2304 VCAs, but by whom or why is uncertain.

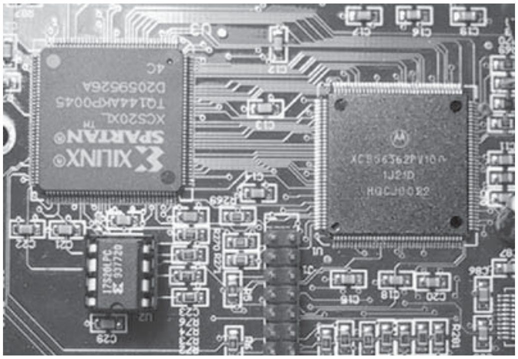


Figure 25-125. Mix stage of a digital console.

Digital recording and transmission—including the most far-reaching domestic example, CD's are covered extensively in other sections of this book. That these are where digits first made their mark on pro audio is hardly surprising; once the speed of the associated processing speed and bandwidths were high enough, well-proven techniques from the communications and computer world were applied to the problems of storing and moving the fairly prodigious number of bits digital audio demands. After all, the major telephone companies worldwide have been using high-speed digital streams for decades. Early successes for audio include the conversion of the BBC's nationwide radio network program distribution system to digital in 1971. The turn of the 80s saw the first few serious digital tape machines, heralded by the 3M/BBC design; then a little thing called the PC happened. Hard disk recording moved from the high-end esoteric to the bedroom studio and is now both ubiquitous and universal. And very good. The pro-audio digital revolution is almost complete. Resistance is futile.

25.17.1 Digital Audio Systems

Fig. 25-126 shows about the simplest example of a DSP (digital signal processing) system possible. The processor itself, in old days racks of discrete logic and latterly specifically tailored microprocessors, is sandwiched between means of coupling it to an analog world outside. We'll first look at the converters and then the DSP bit.

25.18 converters

25.18.1 A/D Conversion

25.18.1.1 Resolution

A DSP processor needs a stream of digital words of sufficient resolution to adequately portray the actual input signal level at a given moment. This resolution is determined by the number of binary bits in each word; each bit corresponds to a doubling of the resolution, or roughly 6 dB of dynamic range capability. Phone systems typically use 8 bits (approximately 48 dB dynamic range linear, although effectively more when companded), the BBC's original distribution system was a 13 bit system (78 dB), CDs 16 bit (9 dB) and most production and recording systems a nominal 24 bit.

The A/D conversion process is fraught, particularly with high-resolution converters, and the actual dynamic range is often much less than theoretically possible from the number of bits. System noises, either from the analog paths or crosstalk from various digital signals, are the predominant limitation; gross errors used to come from nonmonotonicity. Strictly speaking, a converter should, if the input signal is increased by one unit of the resolution, reflect this by increasing the value of the output digital word by one bit. Often, particularly at transitions of the major bits, this goes awry and an

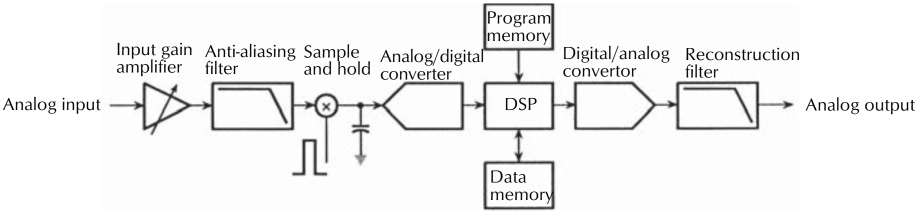


Figure 25-126. A basic digital signal processor in an analog world.

untoward jump in output level occurs. (As an example, the transition from 01111111 to 10000000 in an 8 bit word is a likely point of nonmonotonicity. Although this only reflects one increment of resolution change, a lot of convertor bits are changing simultaneously; the more that have to change—especially the wider the word—the more chance for error. Trust is being laid in the converter's manufacturer that each successive bit carries *exactly* twice the weight of the previous one; with very wide converters the increments of resolution are tiny and the odds are increasingly slim. In a 16 bit converter the most significant bit has to be accurate to within at least a bit of resolution for the device to be monotonic; this corresponds to an accuracy of better than 0.0015%. Enough said. The more bits change in a transition, each of the individual bit's tolerances come into play and errors are far more likely.

The almost wholesale shift to sigma-delta–type A/D converters, which are inherently monotonic, has all but buried this problem now in most practical circumstances. Integrated IC sigma-delta converters are available very inexpensively with what used to be considered science-fiction performance. As mentioned, the primary limitation is noise, either induced digital mush or from the necessary analog parts in the mixed-signal format of these devices; such is not an ideal environment for low-level analog. The low supply rail voltages (5 V is considered big these days) mean that the additional dynamic range available from conventional high rails (such as typically ±15 V or more) is simply not available.

Although it is actually quite difficult now to find a convertor that is rated at any less than 24 bits (and to be fair, their internal structure, in particular the word width of the FIRs, is 24 bit), the actual performance depends on how many of those bits represent useful data and how many are marketing bits.

25.18.1.2 Sampling Rate

In addition to the required resolution, speed of conversion plays a great part. In order to give an accurate portrayal over time of an input signals waveshape, there need to be enough conversions for the digitized signal to be reconstructed to an exact analog of the original signal. The lowest theoretical (Nyquist) sampling rate is twice the highest frequency intended to be processed. This implies at least two digital word conversions taking place for each cycle of (typically in audio) 20 kHz. In practice the sampling rate is made even higher, and figures of 44.1 kHz (domestic) and 48 kHz in professional audio are the most common, 96 kHz and higher yet looming and in search of a mainstream application.

25.18.1.3 Convertor Limitations and Requirements

Currently the figures of 24 bit linear conversion at a 48 kHz rate are de facto standard values in pro audio. Although these parameters are capable of very respectable sonic performance, certainly comparable or in excess of the analog recording and transmission methods digital has supplanted, their practical implementations fall somewhat short of the performance of analog electronics. This is not a snipe at digital; it is clear to anyone who chooses to investigate that the practical differences are small.

The first question is of how much resolution is actually required. A good quiet balanced-bus multitrack console's typical input-output path can be expected to have some 26 dB head room above an operating level of 0 dBu and a noise floor some 90 dB below that for a 116 dB dynamic range. A similar quality mixer summing a fair number of sources can still be reasonably expected to have a noise floor of –80 dBu corresponding to 106 dB dynamic range. These values imply digital word widths of some 20 bits and 18 bits, respectively. converters of these capabilities are readily available commercially, if implemented well. A further related question is what is the highest dynamic range signal source? A very good condenser microphone with a very quiet FET in it in a very quiet room is probably the best candidate; it might be able to cope with a 130 dB SPL gunshot at the high end of the range while still hearing breathing noises in the room at the other. Why one might want, other than as a science experiment, to encompass this total range without gain adjustment begs to be answered. Practically, though, the dynamic range of nearly any meaningful source (that could end up in a mix) or a finished product (that people might want to spend money to listen to) is actually considerably less than that of available digital hardware implemented and used properly. However, one must not get too cavalier with this approach—clean masters (either multitrack or mixed) should retain a dynamic range well in excess of the intended distribution medium to allow for losses in reduction and processing of the masters for same. Just a little compression can put a big dent in dynamic range.

Second is a matter of sampling rate. This, possibly because it is of less obvious sonic consequence than resolution, effects due to its have taken longer to appreciate and resolve. Although Nyquist sampling can, no question, allow the reconstruction of near half-Nyquist frequencies, it belies the fact that two samples per cycle is insufficient to determine the frequency of and the exact level of a signal if either is changing—a considerable number of samples are needed over several cycles and even then the dynamic reconstruction is pretty smeared and sloppy as a consequence of the time-domain response of the reconstruction filter. Nyquist sampling works quite well for audio since there isn't that much sonic information up around 24 kHz, and by and large what there is doesn't move too quickly. But this smearing effect of close-in band-limiting antialiasing filters really made a mess of highly transient material in those big bad days before oversampling.

25.18.1.4 AntiAliasing Filters

Early converters used brutal brick-wall filters at the Nyquist frequency to prevent ultrasonic frequencies from being mirrored down (aliased) into the audio and prevent further ultrasonic signals from heterodyning with the sampling frequency.

For example, a 40 kHz signal passing into a 48 kHz rate sampler will produce an 8 kHz by-product that will definitely become audible upon signal reconstruction. There is no way these filters could ever be described as anything other than a bad thing. Their temporal response was appalling, their effect reaching far, far down into the desired audio passband. More than anything else, it was these filters that gave digital audio a bad name in its early days.

Sigma-delta converters come to the rescue. Oversampling, a technique of taking and reconstructing samples at a multiple of the sampling rate (4, 8, 16, or so), allowing the nasty filters to be both relaxed in brutality and moved correspondingly higher in frequency, dramatically improved this situation—the filters had far less in-band effect. Sigma-delta converters typically initially sample 64, 128, or even 256 times above the nominal sample rate with the consequence that the antialiasing can be reduced to as little as a gentle single or double-pole filter; the band limitation is done inside the converters by a phase-linear FIR filter, with considerably reduced sonic impact.

Nevertheless, there have been experiments that indicate that even such benign internal filters at 20 kHz are with some program material and under some conditions audible, in comparison to the same class of filter set twice as high in frequency. Since the only way to properly engineer such a filter at 40 kHz is to double the sample rate, it seems that the predominant improvement (and ever-so-slight at that) of a 96+ kHz system is not the increased bandwidth available—arguments will continue to rage about our ability to hear/sense stuff up there, and even the desirability of its existence—but that doubling the rate is the only means of pushing the last vestiges of filter effects from audibility. Since this means doubling the amount of processing hardware in a system, it is not a light decision.

25.18.1.5 Types of A/D converters

There are three types of converters with possible application to digital audio. Although without question the sigma-delta type rules the roost in pro audio, enough applications use flash and successive approximation for them to be considered here.

Flash Conversion. Flash conversion involves a long train of comparators, such that a given signal amplitude will trip a given number of comparators and fairly simple conversion logic can turn their outputs into a binary word. It is the fastest conversion method as far as logic propagation times; a change in input level is instantly reflected in output code. The down side is the sheer number of comparators needed for a sensible size word width, one for each possible level of resolution; also the offset inaccuracies of the comparators tend to dwarf the required resolution! This said, they are little used, except in some hybrid converters where a 4-bit flash convertor will provide the major resolution of a wider word, leaving the remainder to a more accurate type.

Successive-Approximation Encoder. The successive-approximation encoder is a very common form of encoder, especially where high speed at high accuracy and with low latency (processing delay time) are required. There is but one comparator unlike in flash, easing accuracy. Conversion takes at least as many cycles as there are bits of resolution. Operation consists of comparing internal voltages, weighted in accordance with the bit value, against a frozen (by a sample and hold circuit) sample of the input signal. It needs to be frozen since the conversion is not instantaneous and the input signal level could change in the time it does take. The most significant bit's value is half the permissible input range, the second a quarter, the third an eighth, the fourth a sixteenth, and so on in binary weighting. They are applied in turn to the comparator, MSB first. If the sample is larger than the MSB, then the MSB is left asserted; if not, it is dropped. The next weight is applied; if the input sample is still larger than the combination of MSB and No. 2, then No. 2 is left asserted, and so on. Eventually all the bits are tried against the input sample with the bits remaining asserted, forming the 1s of the digital word, the remainder the 0s.

Both of the above converters generate an absolute digital value of the input signal at each sample period.

Sigma-Delta Conversion. Sigma-delta, or also called delta-sigma, conversion starts off in essence by measuring relatively how far the input signal moves, up or down, rather than stating exactly *where* it is. Conversion occurs much more often than the required output sampling rate (e.g., 48 kHz) often 128 or 256 times higher. The conversion itself is much simpler though. Simplistically, at each conversion it only has to make the decision whether the input signal has moved up or down from where it was last sampled. Its output is a very fast stream of up and down signals; the sampling is fast enough that it can keep pace with the input signal's probable changes, sensing automatically whether large level shifts or tiny ones are taking place. Subsequent intelligence (filtering) keeps track of this torrent of single-bit state changes and renders down a conventional digital word for an absolute output value.

As a method it has many advantages, not the least of which being the enormous internal sampling rate; the antialiasing filter can be relaxed considerably, both in order and cutoff frequency (often it just consists of a single- or double-order filter set much higher in frequency than with other encoders—and sometimes left out completely!). Filtering is left to within the digital domain.

They are also monotonic, having none of the problems of the other types of comparator level or ladder accuracy. What they do have, and which can be a concern in some applications, is a comparatively very long latency (signal processing delay time) before a relevant sample pops out for digital digestion; at normal sample rates and depending on the length of the FIR decimation filters within, this latency can be around a millisecond or so. Sigma-delta A/Ds predominate in pro-audio.

25.18.2 Digital-to-Analog Conversion (DAC)

25.18.2.1 Conventional Ladder DACs

A means of turning the processed output signal from the DSP back into analog is necessary. These are described in Chapters 31, 38, and 39, but for completeness are outlined here. A DAC adds together voltages (or currents) of weightings corresponding to the importance of the binary bits. Fig. 25-127 shows a simplistic DAC. The required output digital word is applied and the most significant bit, if set high, sources a current of 1 mA. The next most significant bit sources half that or 0.5 mA, the next bit half that (0.25 mA), and so on down to 7.8 μA increment for the least significant bit. In the 8 bit converter shown, the maximum output current is just short of 2 mA (1.996 mA) with all the bits set (one extreme) and none if all are low (the other extreme). Any current between those two, in 255 steps, which is the resolution of an 8 bit word, can be achieved by setting up a permutation of the input bits. This output current can be converted to an output voltage by a summing amplifier.

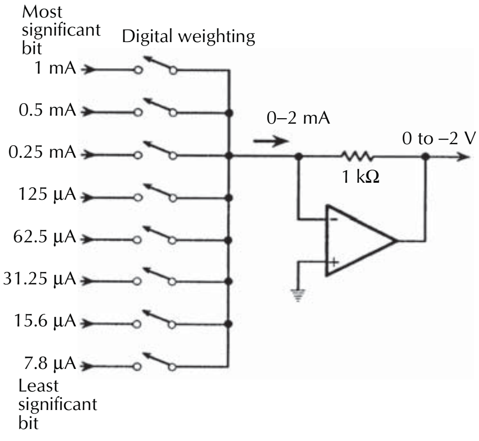


Figure 25-127. Simplistic digital to analog converter.

There are other kinds of D/A techniques, probably the most common being the R/2R ladder, Fig. 25-128.

Much as the simple DAC, asserting a bit causes a correspondingly binary weighted current to be output.

25.18.2.2 Reconstruction Filters

All the earlier comments about antialiasing filters apply here, too. As well as the required audio—up to 20 kHz in bandwidth—coming out of the DAC there are a host of other products, the most unappealing and closest in spectral terms being a mirror image of the audio centered on the sampling frequency and descending in frequency; a 20 kHz audio signal sampled at 48 kHz will be output from the DAC along with an image at 28 kHz (sample frequency minus audio). Heterodyning strikes again; there will also be an image at 68 kHz (sample frequency plus audio), and in all likelihood more sets of images centered on harmonics of the sampling frequency. The most dangerous sonically though is that first inverse image.

25.18.2.3 Oversampling

Enter a filter every bit as precipitous as the one needed at the front end. Every bit as nasty, too. Solutions other than good, well-designed filters come from the digital domain; oversampling, for one. One approach is to intersperse an *interpolation filter* between the processor and DAC. This digital filter reconstructs the audio but at a higher sample rate; the smoothing of the filter effectively creates more sample points between the few actually being issued by the digital source (DSP). If a guessed digital word is inserted between each of the real ones, the effective sampling rate becomes doubled and, in practice, the DAC is working twice as hard and fast outputting analog.

Here is the good part. If the sampling rate is doubled, the heterodyning images start that much higher up in frequency; following the earlier example through, a 20 kHz signal's first inverse image is now going to be at 76 kHz (96 kHz minus 20 kHz) instead of 28 kHz as before. The immediate benefit is in the relaxation of the reconstruction filter—it can be much less steep and pushed up in frequency somewhat away from the audio band.

The oversampling process can be carried on even further; four times, eight times, even sixteen times and more with greater oversampling rates commonly used, pushing the undesired products correspondingly higher in frequency and so dramatically relaxing reconstruction filter requirements. The fact that fifteen out of sixteen samples may be filter guesses belies the fact that it isn't those that improve the audible performance—it is the absence of brutal analog filtering that makes all the difference. Exactly the same conditions apply here to the application of higher (96+ kHz) sample rates, simply with the intent of pushing antialiasing filter effects out of audibility, as in A/D converters.

25.18.2.4 Sigma-Delta DACs

Sigma-delta DACs oversample to the same degree (64, 256, or beyond) as their previously described A/D brothers and the corresponding increase in frequency of the reconstruction filter dramatically simplifies their implementation. Most D/As in proaudio are now sigma-deltas, although conventional laddertypes are still in wide use, and especially where higher than normal audio speed is required (such as in a broadcast stereo encoder). Again, latency is the only major drawback to this type; the processing delay again depends on sample rate and the particular device and its filter length, but is generally around a millisecond. This, of course, means that a system using sigma-deltas at both ends (ADC and DAC) can potentially have a latency of a couple of milliseconds or so; this can be a bust in some applications.

25.18.3 Sample-Rate converters (SRCs)

A big problem facing digital audio system designers in the early days was combining sources from different machinery that, unless heroics were performed and all the system's machinery was phase and word-clock synchronized, almost certainly were all running at

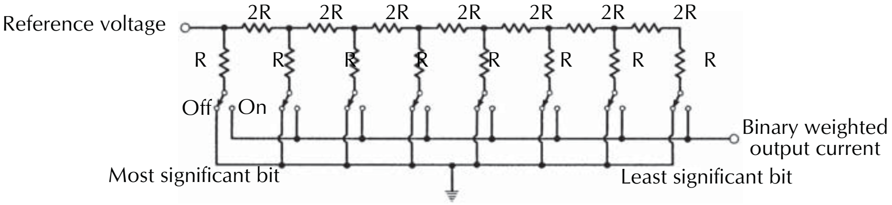


Figure 25-128. An R/2R digital to analog conversion ladder.

ever-so-slightly-exactly-not-quite on the same frequencies, from their own independent clocks. Mixing such is a disaster.

SRCs allow sources with a wide range of sample rates to be reclocked to the console's master clock, allowing them to be processed normally. Being that they use internally very long FIR filters and of varying length depending on the ratio of input-to-output rate conversion, they not only have latency, but it changes too. Another slight shortcoming is the tendency to affect the ultimate dynamic range by leaving artifacts way down in level, but most current parts are excellent in this regard. All in all, they are a near miracle-cure for what was an intractable problem.

25.19 Digital Signal Processors

There are a number of features that distinguish devices specifically designed for DSP from the admittedly bigger and faster but generally dumber and seriously more expensive behemoths powering PCs and the like.

25.19.1 Multiplier/Accumulator (MAC)

The heart of a DSP device is the hardware multiplier in its arithmetic logic unit. This takes two full data-width numbers, multiplies them, and leaves the result in an accumulator, quickly. Further products of multiplications can be arranged by a software instruction called a MAC, Multiply/ACcumulate, to be added to results previously stored in the accumulator. The MAC is central to DSP. Nearly every manipulation of a signal in the digital domain is ultimately achieved by multiplying a sample by another value, called the coefficient. The simplest example is that of level control—in audio terms, gain control. If an incoming sample is multiplied by a value of 1, the result lodged in the accumulator is the same as the input sample. If the gain-defining coefficient is greater or less than 1, the accumulated result is correspondingly greater or smaller than the input sample.

The accumulator necessarily needs to be of a wider word width than the input byte width capability since a multiplication can end up with a much bigger or smaller number than the input sample; in the case of the very popular fixed-point Freescale (see Motorola) 56300 series DSP chips, the bus width (input-output word width) is 24 bits while the accumulator widths are 56 bits. It's a worthwhile rule of thumb—a multiplication results in double the bit width.

In this volume-control example, the input analog signal is sampled at the front end of the encoder, an A/D conversion is performed, and this value is deposited on the DSP chip bus at its command. The input word is multiplied by a coefficient, similarly picked up off the data bus, and the result left in the accumulator. Rounding off fits the possibly too long result to the width of the DAC (e.g., down to 16 bits from a possible maximum result of 32 bits from one 16 bit multiply). The answer is put on the bus to be picked off at command by the D/A convertor. The D/A performs a near-instantaneous conversion back to analog, ready for consumption by the real world. This whole routine is repeated 48,000 times a second; each operation has less than about 20 μs to take place. Congratulations! This is the digital replacement for a $5 potentiometer.

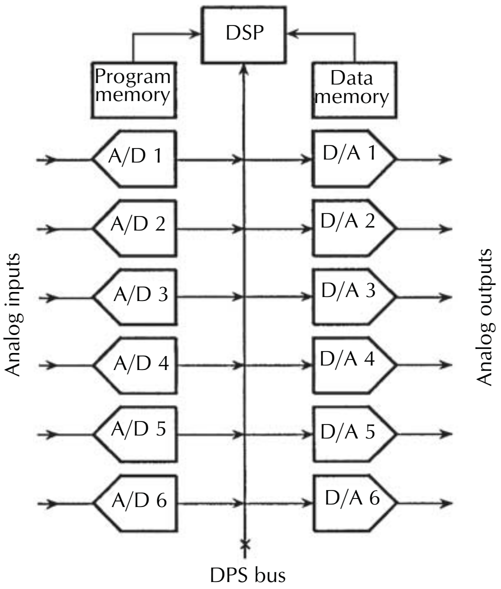
To get a sense of the great strength of the digital solution, Fig. 25-129A shows many A/D and D/A converters hanging on the DSP input-output bus. Each of these is independently addressable by the DSP chip; it can systematically pick an input signal word from any A/D, work on it, then deposit the result into any D/A converter. Further, it can take input samples from any or all of the A/Ds, multiply them in differing degrees according to differing coefficients, and add the results progressively within the accumulator. This accumulated result is then scaled and passed to a D/A. In effect, this is the digital equivalent of mixing a number of sources, all the sources at different gain settings, to one output.

The comparatively simple digital arrangement can be made to equate to an analog soft matrix, as drawn in Fig. 25-129B. It's starting to look more like a viable cost and space saving replacement; this small example of six-in and six-out is already equivalent to 36 VCAs.

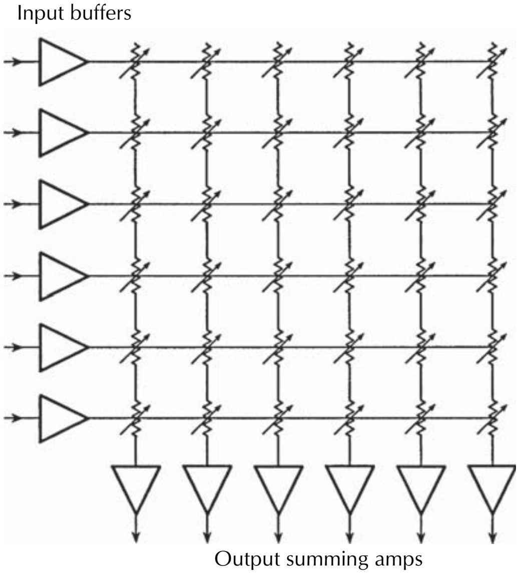
More inputs and more outputs to and from the mix stage are of course possible. The principal limitations are accumulator width, which is taken care of by building in adequate head room just as one does in analog, but more importantly processor time; after all, it still has to do all the input-to-output multiplies within a 20.8 μs window.

25.19.2 Instruction Cycles

A processor instruction cycle is simplistically the time it takes to perform one single simple operation, such as a bus access (to acquire or dispose of data), an arithmetic function or a move of data from a register to elsewhere. Multiplies can take a bit longer, depending on the chip, but DSP chip architectures with hardware multipliers are very slick and time efficient. They need to be. The processor speed determines how many of these clock cycles are available for processing in a given time



A. Multiple level controls using DSP.



B. Analog equivalent.

Figure 25-129. A digital mix of a number of sources.

window and so directly limits how much work the processor can do. For example, a 200 Meg device has a processor cycle rate of 200,000,000 Hz. Given a 48 kHz sampling rate, this gives a maximum of just over 4000 cycles per sample period. Some operations can take more than just one clock cycle, so this is an outside ideal figure. In practice, it works out at somewhat fewer. Although it looks like a big number, it seems vanishingly small as soon as anything clever is attempted with the DSP. This, above all else, is the primary reason why upping the sample rate above the bare possible minimum is a very unpopular notion in DSP circles. Cycle budgets rule.

25.19.3 Processor Types

Specific DSP devices are chosen for a wide variety of reasons, both real and perceived. Device flexibility, per-unit cost, and ease of implementation (in the forms of support from the manufacturer and quality of the design tools) all factor in. In very large run products such as consumer items, part cost will probably override everything else while ease and speed of implementation tend to be more important in lower-volume, high-tweak-factor arenas such as pro-audio. Rarely is there one overweaning performance feature that makes or breaks a choice. However, since in order to squeeze the most processing from each device a considerable amount of their programming is still at the machine-code level, the designer's familiarity with a particular assembler language can have a strong influence—this definitely falls into the ease and speed of implantation department.

Perhaps the minimum for processing audio data is a 24 bit word width and correspondingly wider accumulators and registers. As such the Freescale devices just about fit the bill. They are *fixed point* processors, which directly limits their dynamic range to the number of bits (144 dB for 24 bits, 336 dB for the accumulators); fortunately, this is plenty for most real-world audio processing. Some applications, like some filters, demand wider immediate dynamic ranges in their calculation and intermediate-value data storage, and for those instances long or double-precision arithmetic is used. The down side is that such filters can take up to twice as long (twice as many cycles) to calculate as single precision.

25.19.4 Floating Point

Floating point processors (*floaters*) as exemplified by Analog Device's "Sharc" series avoid this problem by representing numbers internally in exponent/mantissa format, having far more involved internal processing to handle the complexity of dealing with these numbers. The Sharcs can be operated as either 32 bit fixed point or 32 bit floating point. Since the dynamic range of a floater is as good as infinite regardless, none of the dancing around one sometimes has to do with a fixed point applies. On the other hand, the capacity to dig a big hole is as good as infinite, too.

25.19.5 Parallelism

DSPs differ from conventional microprocessors in that their architecture is contrived to make certain common processes as slick as possible and to be able to perform as much real data manipulation and housekeeping within each clock cycle as possible. This latter is called *parallelism*, and the degree of parallelism is what sets devices apart in capability. For instance, performing an FIR filter (or a mixer routine, for that matter) within one clock cycle, a DSP can.

- Do a multiply and accumulate (MAC).
- Fetch in the next data word ready for the next MAC, update data pointer.
- Fetch in the next coefficient ready for the next MAC, update coefficient pointer.
- Update the program counter.

In short, everything that needs to happen to ensure a filter point can be calculated within one cycle is done, ready for the next.

25.19.6 Multiple Memory Spaces

Memory allocation is nearly always somewhat different from ordinary processors, which usually have just one memory space shared between all functions; DSPs have at least two separate memory spaces; the Freescales have three, for example: one for the program information (program memory), one typically for coefficients, and one typically for the inevitable intermediate filter values etc. necessarily stored between per-sample calculations and for internally stacked-up audio data, if brought in enbloc from outside.

25.19.7 Real-Time Specific Peripherals

Additionally, most DSPs have convenient peripherals built into them to allow ready, seamless, and fast transfer of data in and out of the chip, either into memory-mapped data space and/or through a variety of serial communication formats. It is usually possible to seamlessly connect a DSP to a number of conventionally serially formatted A/D or D/A converters and to other DSPs. Definitely not least, a ready and fast means of importing fresh programs/coefficients with which to modify the data is always available via a host port.

A related major tool is DMA, or direct memory access. This allows the moving of considerable amounts of data into and out of the DSPs memories with little impact on the main cycle budget other than that required to set up the necessary pointers and to fire off the DMA activity at the required times. On very busy processors conflicts can arise (DMA does borrow some real processor resources while it's not looking, and under normal circumstances usually gets away with it) so it's not entirely free, but it is more than handy.

25.20 Time Processing in DSP

25.20.1 Time Manipulation

Something what is very readily achieved in digital is storing information, either long term onto disks or flash memory, medium term in RAM, or short term within processor registers and internal RAM. Nearly all manipulations of data of any complexity greater than the soft matrix example above demand storage.

25.20.2 Delay

A stream of input data is written into RAM memory and subsequently, some time later, read out again. The length of time recordable (sample length) depends on the size of memory and the sampling rate—the faster the rate, the quicker the memory will be eaten. This memorized sample may be stored elsewhere then— say, on a hard drive.

Say a relatively short time delay is required for an echo. The input data stream would be written into RAM and read out at a fixed time (a certain number of samples) later. Sooner rather than later the memory would run out and the delay stop, so the memory is usually arranged as a circular buffer; when the buffer end is reached, the memory register leaps back to the start of the buffer and overwrites what was previously there, and so on. The buffer is read in the same manner, at a time after it has been written determined by the required delay. As long as the buffer is long enough to contain enough samples for the required delay, a continuous delayed output version of the input is available. The main advantage to the seeming complexity of the circular buffer is that only pointers, or indexes, are being changed and updated; the only audio samples being changed are the writing of the newest sample over the oldest one. What is important is what is *not* happening; huge amounts of data are not being read and rewritten somewhere else. The complexity is merely keeping track of those read and write pointers, which is in reality simple arithmetic and indeed an automated function in many processors.

25.20.3 Echo, Echo, Echo

Reentrant, or recursive, delay (spin echo) where a delayed signal repeats continuously until fading away is achieved by attenuating the delayed words (in the multiplier) and adding them in the accumulator with the concurrent new input words.

If the delay is made very short, and the delay is summed in the accumulator with the new, direct sample, something interesting happens—a direct parallel with the analog world. The direct and delayed signals sum and interfere. A 1 ms delay corresponds to a half wavelength of 500 Hz; in other words at 500 Hz with 1 ms delay, the delayed signal will be out of phase by 180° from the input signal. They will cancel and a notch at 500 Hz (and every 500 Hz interval up the spectrum) will occur. Altering the delay time alters the frequencies at which cancellation occurs; studio people call it *flanging*; we'll call it a comb filter, our very first digital filter.

25.20.4 Reverberation

In real acoustic environments, reverberation is the summation of countless random time-delayed reflections and rereflections from floors, walls, ceilings, and obstacles. Complications set in with differing reflections having differing frequency aberrations due to varying surface absorption coefficients, but in essence it is an accumulation of time-delayed signals of various and decreasing levels. As such it can be reasonably well emulated in DSP by more or less complex variations on time delay; relatively long time delay loops are established to emulate major room reflective modes. Many short loops and all-pass configurations are used to emulate the decorrelation that occurs in an acoustic space by multiple short reflections and diffraction. The output-to-input feedback—terms for each of these elements is *adjustable* and *equalization*, typically in the form of simple roll-offs—are applied either after a loop or within its feedback path to mimic the typically higher absorption at higher frequencies in an acoustic environment. There are a lot of small and large elements, all with a lot of handles, or things that need to be fed parameters.

Basically, the number of elements and the skill in determining their convoluted interaction and parameters decide how convincing the reverberant effect is and its characteristics. Some astonishingly good results have been had from DSPs with quite small (64 k word) external memories.

As DSPs become more powerful and much cheaper, becoming increasingly practical is a class of reverberation units that in effect perform a very, very long convolution of an applied audio signal with a digital recording of the reverberation tail of a real venue (see Section 25.21.1, Transversal Filters, for the basic technique). This can involve hundreds of thousands of DSP multiplies (meaning *lots and lots* of DSPs) but is as one would expect highly impressive and flexible. Proprietary convolution algorithms can reduce the computational burden, but it is still nevertheless a big proposition.

25.20.5 Averaging

An average of a number of input samples is achieved by adding all the input word values for the period of time over which the average is required; this is normally figured out by numbers of samples—20 ms worth of samples at a 48 kHz rate equates to 960 samples. (This would be a l-o-n-g train of samples.) These samples are all added in the accumulator and then divided by the number of samples—the result is an average value for that 20 ms. If each sample is stored elsewhere, then a rolling average becomes possible; for each new input sample added in, the first sample of the 960 is subtracted, and a new average for that instant is calculated.

Division, as such, is something undertaken only under extreme duress in DSP; it is very thirsty and inefficient. A division in such a case as creating an average as here could be achieved by first arranging, if possible, that the average length is a binary interval (2, 4, 8, 16, etc.). Then the end result of all the additions could be bit-shifted right the corresponding number of times. An arithmetic shift right (moving a digital word one step to the right, filling in the now missing top bit with a zero) is the same as dividing by two; an average of 64 samples would thus need six right shifts. Alternatively, a single multiply by 0.015625 (1/64) (or the reciprocal of whatever the arbitrary number of samples may be) does the job. Either is an awful sight quicker to do in a DSP than a 24 bit division. *Anything* rather than divide.

25.21 DSP Filtering and Equalization

25.21.1 Transversal, Blumlein, or FIR Filters

Yes, Alan Blumlein invented these, too. The train of samples concept becomes very valuable in DSP. This type of filter can produce a wide variety of time effects and frequency response shapes, particularly bandpass and cutoffs. While the determination of coefficients for the various filter types is beyond the scope and intent of this section, the underlying principle is shown in Fig. 25-130. For each sample period (i.e., every 20 µs) a fresh input sample is inserted at the head of the train; all the samples move along the train and the oldest one falls out of the other end and is lost. Each sample is multiplied by a coefficient specific to that position and summed in the accumulator with other results from the other multiplied samples. Each pickoff is subject to a different coefficient and sum sense (normal or inverse). The accumulator value is the new output word for that particular sample time; 20 µs later the whole routine starts over again.

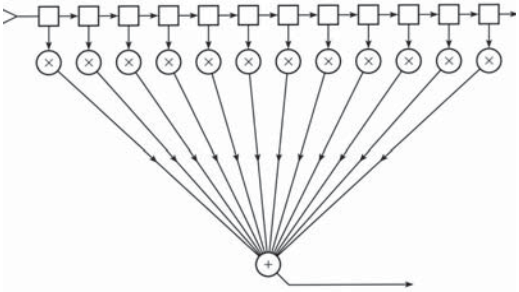


Figure 25-130. Transversal or finite impulse response filter (FIR).

This passing of one set of data (in this case audio) *through* another set of data (coefficients) is also called convolution.

25.21.1.1 Impulse Response

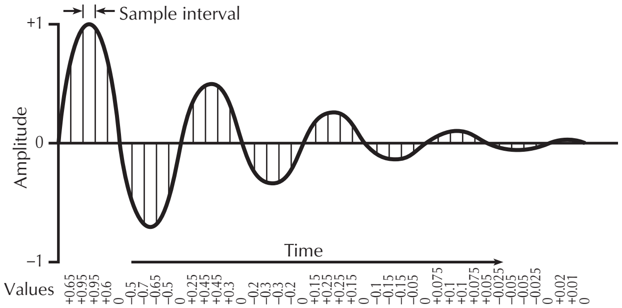
As familiar as we are with using the tool of frequency response measurement to analyze or describe the transfer function of a device or circuit, there is an equally powerful descriptor: the impulse response. Embracing the impulse response concept aids gaining a mental picture of how digital filters work. Fig. 25-131A is what the waveform of a large bell excited by an impulse could look like: a damped sine wave at the tone of the bell. (Hardly dissimilar to that from a damped oscillator, or bandpass filter. Hold that thought in mind. Actually, looking at the response, it would probably sound more like the dung of a lamp post, but please suspend disbelief for now.)

Each of the vertical lines represents the instantaneous amplitude of the signal at each sampling period; this, if you like, is a sample-by-sample digital recording of the bell's sound. If we were to play back the samples at the rate they were recorded, we would hear the bell thunk again. We now use the bell's samples' numeric values as coefficients in a transversal filter, Fig. 25-131B, and send an impulse (one sample of full positive amplitude, the rest zero) into the filter; the effect is exactly the same. As the impulse passes each coefficient, the bell sound will be reconstructed once more.

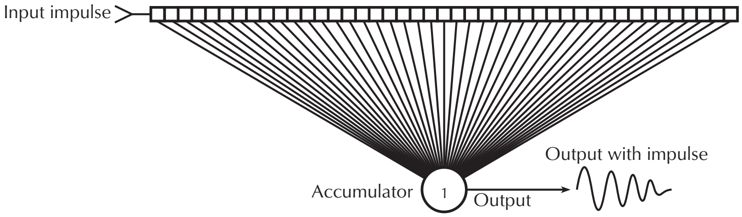
There is nothing to stop us from putting real audio samples into the front of the transversal filter—the effect will be as if the audio is being played through a damped bandpass filter at the frequency of the bell. The bell's impulse response is impinging itself directly on the audio passing through the transversal stages. It will sound as though you're listening to the audio with your head stuck up inside that bell. Yes, it's a filter! In short, if we can describe a desired filter's impulse response and use its samples as coefficients in a transversal filter, any signal passing through the transversal stages will be filtered accordingly.

This kind of processing is commonly called FIR (finite impulse response) filtering. If a transient (impulse) were encoded and applied to such a filter, the samples describing it would enter the train of stages. Output summation contributions occur until they reach the end. When the last relevant sample has fallen out the end of the train, no further output samples that have anything to do with the originally applied transient are possible. The duration of the transient within the filter is limited to the lifetime of its samples in the train; they eventually all leave. The impulse's existence is finite. The filter's length is finite—hence, finite impulse response.

Intellectually, FIRs are very appealing through their very simplicity. Unfortunately, this genre of filtering is rather taxing in current DSP terms since it demands a lot of processor time for any useful audio filters. As a rough rule of thumb, to do anything meaningful at a given frequency the filter must be able to contain a full cycle of that frequency; to operate at 50 Hz an FIR would need to be at least 20 ms long, which (assuming a 48 kHz sampling rate) would be about 1000 filter points long. As mentioned earlier, a 200 MHz part only has about 4000 cycles of processing available per sample—this one filter has just eaten about a quarter of



A. An impulse response.



B. Impulse response used as coefficients in a transversal filter.

Figure 25-131. An impulse response becomes a filter.

a whole DSP! Suddenly, except for a few rather special and esoteric circumstances, such as phase-linear EQ and auto adaptivity, it becomes obvious why FIRs are not particularly popular in mainstream audio DSP processing. They are rather hardware and time thirsty.

Impulse response coefficient sets suitable for plugging into transversal filters may be either calculated (long-handed for the rigorously inclined, or within any of the many excellent filter design programs available) or, as in the only half-joking bell example above, recorded by issuing an impulse into a pet filter and using the resulting sampled output as coefficients—audio played through an FIR with those coefficients will sound just as if it was passing through the original filter. As earlier mentioned, there are reverberation units working exactly on that principle.

25.21.1.2 Windowing

Any attempt to generate a set of coefficients for FIRs will run into the problem that an ideal filter simply will not fit into the length of any practical filter. Obviously, the filter has to be long enough to realistically encompass the meat of the desired processing (a 99 point filter won't do 50 Hz, remember?), but this still leaves the problem that the filter is finite in length. A series of FIR filter designs showing the impulse responses and corresponding frequency responses of a 33 point (33 step long) nominally 12 kHz high-pass filter highlight the quart/pint-pot tradeoffs. Truncation—i.e., lopping the end(s) off to make it fit—leads to Gibb's phenomenon, in which the desired output frequency response of the filter is seriously compromised by large lobes, Fig. 25-132.

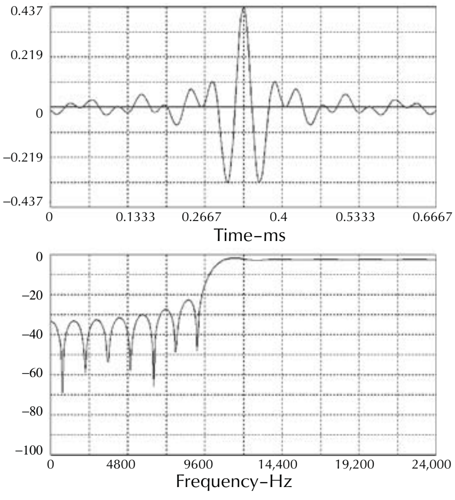


Figure 25-132. A 33-point unwindowed FIR filter.

Mr. Hanning, Mr. Hamming, and Mr. Harris (among others) come to the rescue here, with a technique called *windowing*. These apply weighting to the values of the coefficient set, basically leaving the most significant elements (usually in the middle of the set) alone and tapering off the values toward the ends of the set. The taper that is applied varies according to the type of window, and the differing types are best suited to differing interests of compromise. Say a brick-wall filter had been described as in the figures; one window may optimize for stop-band rejection, Fig. 25-134, another may trade that against sharpness of the filter cutoff rate, Fig. 25-133, etc. Many thanks to Momentum Data Systems' software for the curves.

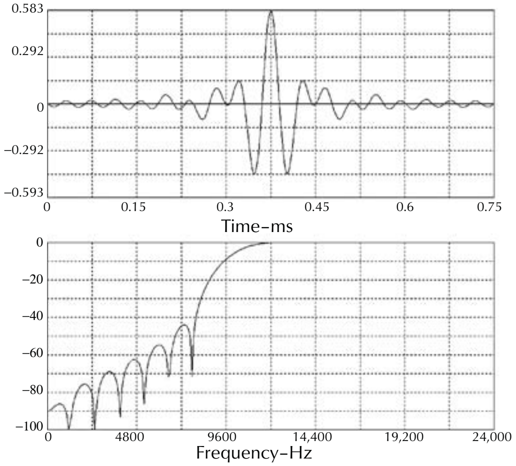


Figure 25-133. The same 33 point filter Hanning windowed.

25.21.1.3 Symmetrical FIRs

There is an FIR implementation that has some quite interesting properties and as such is probably the most used, so much so that the majority of commercial design packages assumes as a default that one wishes to design symmetrical FIRs and that FIR has become almost synonymous with the symmetrical filters they afford.

These allow the imposition of a frequency response (in the case of a conventional-style EQ) *without* altering the phase response, unlike ordinary EQ (and nature) in which any frequency response change comes with a corresponding shift in phase response for free. Although this characteristic might at first blush seem ideal and a major leap forward for audio technology, in practice they are only rarely used; yes, Virginia, they do sound different to conventional EQ with equivalent frequency responses, but not necessarily better. (An odd effect is that one seems to need more phaseless EQ cranked in than conventional EQ for a similar subjective effect.) Certainly, it's not better enough to displace conventional EQ, which can be readily and far more efficiently

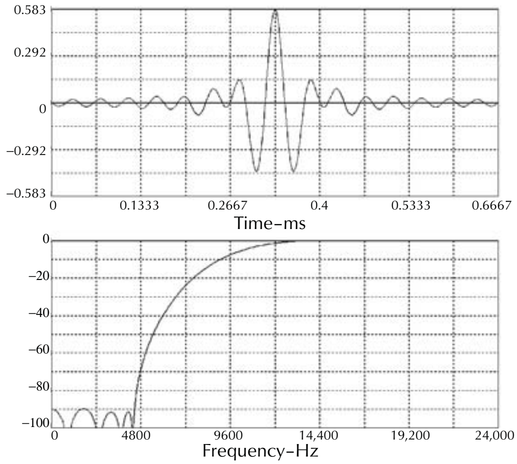


Figure 25-134. 33-point filter Harris windowed.

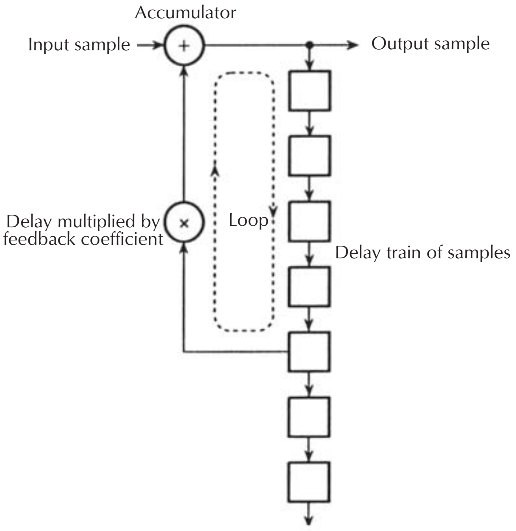
created in either digital or analog form. The difference alone, however, is sufficient reason for existence in music production, and special-effects units and audio workstation plug-in software specifically to do symmetrical FIRs are available.

Symmetry refers to the fact that the coefficient set is arranged to be symmetrical about the center—the midpoint—of the filter; identical coefficient set-lets tail off toward the end of the set as tail back toward the front. The midpoint of the filter is regarded as the time center—in other words, a symmetrical FIR has an intrinsic time delay of a passed signal of half the length of time the filter takes to calculate; in our now-famous 50 Hz capable, 960 point filter, the effective time delay is 10 ms, or half of the time it takes for any one data sample to transit the entire filter, being 20 ms. This time delay is another major downside to symmetrical FIRs; in order to keep everything in a multisource console time aligned, all other sources would have to be delayed by the effective time delay of just one FIR'ed source.

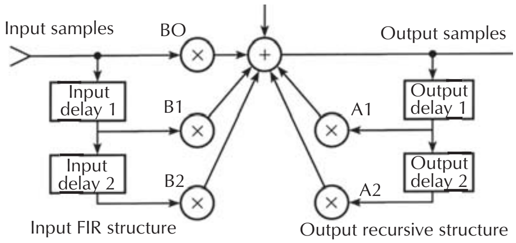
Note that not only is half of the filtering done after to the time center, but, and this is the head hurter, half of it is done *before* the time center, leading up to it. The filter only remains causal because of the intrinsic time delay. That the ear can deal with filtering effects before something has happened and integrate it all into an acceptable sound is a true amazement.

25.21.2 Recursive Processing

This concept was approached in achieving spin echo and reverberation; feeding an already manipulated input sample back around in a loop to be reprocessed along with new, and/or yet other samples. Fig. 25-135A shows this diagrammatically. A time delay (a number of samples' delay) is included in a loop and fed back at a level determined by the controlling coefficient. Picking off different samples and treating them with differing coefficients allows great control over the nature of the feedback and the dynamic nature of the loop. The most important thing to note is that once a signal has entered the loop, it just carries on going around and around, being summed with fresh input samples each time. The time taken for a signal to die away is determined by the coefficients in the feedback loop—this can very loosely be paralleled to the analog concept of *Q*; the more positive feedback in a filter the tighter its response, with the drawback that ill attention to its control can result in oscillation. Such is exactly the case with digital recursive stages. Even if controlled, the signal never actually dies away completely; in DSP this can result in leftover bits rattling around, manifesting as repetitive cyclical errors. Sufficient accumulator width needs to be available to round or noise-shape results off nicely.



A. Simple recursive processing.



B. Direct-form 1 biquad IIR filter.

Figure 25-135. Recursive Processing.

The first big advantage of recursive processing is that significantly less memory accesses are needed than with FIR—history (equating to length of the filter and its temporal resolution) is built up within the loop rather than being necessary individually and sequentially. The second advantage is that far fewer coefficients and operations are needed.

25.21.2.1 IIR Filters

A filter built up around recursive techniques is known as an *infinite impulse response filter* or IIR, Fig. 25-135B, so called because once in the loop, an impulse just keeps trundling around indefinitely, infinitely. In practice it gets rounded off sooner or later, but it makes the distinction from the FIR.

Additionally, an output appears from an IIR at the same time as input samples are applied and the filter starts behaving as a filter immediately; the only delay is the group delay of the filter, exactly as in analog; there's no waiting for sufficient data to be affected by sufficient coefficients for the nature of the filter to become formed, as occurs in symmetrical FIRs.

IIRs are presently easier, quicker, and carry less time and memory overhead than FIR filters; consequently, they tend to be much more popular for audio DSP.

25.21.2.2 The Biquad

There are many different ways of implementing IIRs; the one in Fig. 25-135B is known to its chums as *direct form 1 biquad* and serves to illustrate the process well. (Others may use less memory space, or run a bit quicker, but they come up with the same results.) It's known as a biquad since it essentially calculates a biquadratic equation. There is no real limit to the number of input and output delays, multiplie, and summations; it's just that those longer than a classic biquad tend to be rather high-strung creatures that are far less easy to calculate coefficients for and to keep tame (read stable).

There are two halves to the filter: the input stage consists of a short, 3 tap, FIR. The applied input signal is multiplied by a coefficient (b0) and the result put in the accumulator; the two delay line outputs are multiplied by their coefficients (b1 and b2) and added to the accumulator. The output stage is a two-delay recursive section; each of the two delay-line outputs are multiplied by their respective coefficients (a1 and a2), the results adding to the accumulator, the total contents of which now represent the output. Once used in a sample time's calculation, the contents of the input and output delay lines move down one, such that the value in input delay 2 is displaced by the former contents of input delay 1, which is in turn filled by the last used input data sample; similar actions occur in the output delay, only the last calculated output sample enters the delay line.

In short, the biquad takes all of five coefficients, five MACs (multiply and accumulates) and a bit of concurrent shuffling of data to happily create a second-order high-pass, low-pass, or bandpass filter. They are very quick and easy to implement in DSP. That it is the basic building block of most digital EQ is hardly, therefore, a surprise.

25.21.2.3 Coefficient Analysis

A look at the coefficients for the input 3-point FIR can give a clue as to what class of filter the biquad is running. (Actually, complicated only by the higher number of coefficients, the same sort of analysis can be done to longer FIR coefficient sets, too.) Fig. 25-136 shows three biquad coefficient sets. Set A shows equal and opposite coefficients for b0 and b2, and none for b1. A very low applied audio frequency (dc, or practically so) will present substantially the same input signal level over the three samples in the FIR; this means that what is contributed to the MAC by the input sample by the b0 multiplication is going to be substantially canceled by a similar size signal from b2 (b1 not contributing at all, being zero). So low frequencies are not being passed. By a similar token, a signal of half the sample frequency (Fs/2) would have near identical values in the first and third positions, i.e., being administered by b0 and b2; since these coefficients are inverted this frequency gets nulled too. The highest valid audio frequency (Fs/2) is not being passed, and neither are very low frequencies. With a bit of luck something in the middle will be, so this is in all likelihood a bandpass filter.

On the second coefficient set, Fig. 25-136B, the effect of multiplies by b0 and b2 are entirely canceled by the effect of b1 at dc, yet by virtue of b0 and b2 not being inverted with respect to each other, Fs/2 is not nulled. The conclusion would be, if one didn't already know, that this was likely a high-pass filter.

The only apparently slightly misleading case is that of the lowpass filter, Fig. 25-136C, which one would expect to bung in a null at Fs/2, but doesn't at first glance seem to, being that the b0 and b2 coefficients are the same. Aha! though. If one imagines the positive peaks of an Fs/2 signal being coincident with b0 and b2, then the b1 coefficient is coincident with the negative-going peak—the b1 coefficient, being the sum of b0 and b2, neatly cancels their effect by creating a negative signal equal to their positive contributions at Fs/2.

A.

48 k	bpf	1000 Hz	Q .707
b0		.292543	
b1		0	
b2		–.292543	
a1		–1.98289	
a2		.7074571	

B.

48 k	hpf	1000 Hz	Q .707
b0		.9115751	
b1		–1.82315	
b2		.9115751	
a1		–1.815318	
a2		.8309824	

C.

48k	1pf	1000 Hz	Q .707
b0		3.916071E-03	
b1		7.832143E-03	
b2		3.916071E-03	
a1		–1.815318	
a2		.8309824	

Biquad coefficient sets for bandpass, high-pass, and low-pass filters. Note the similarity in the a1 and a2 coefficients; the class of filter is determined by the b0, b1 and b2 coefficients of the input FIR structure.

Figure 25-136. Biquad coefficient sets.

Fig. 25-137 graphically shows these analysis results. In short the input FIR is a dumb little filter of the same class of the overall filter—and actually is what determines its class—the output feedback IIR structure in effect determining the frequency and *Q*.

25.21.2.4 Filter Quantization Distortion

The output (recursive) stages of a biquad can cause some pretty wild signal levels to be achieved in the accumulator—they are, after all, little more than a slightly complex feedback circuit; the output signal is fed back in part through both delays in accord with their a1 and a2 multiplies and grows until it (hopefully) stabilizes. (This can be likened to operating a PA right on the edge of feedback at microscopically varying degrees; this is standard operating procedure with IIRs). Filters that are either high in *Q* and/or more importantly very low in frequency with respect to the sample rate (and that, unfortunately, means most EQ-type frequencies) exacerbate this effect. The cure is to only excite the filter to the degree that the desired output is unity with respect to the source; this is usually achieved by proportionally scaling back the coefficients in the FIR input chain. Looking at the coefficients in Fig. 25-136C, this can be

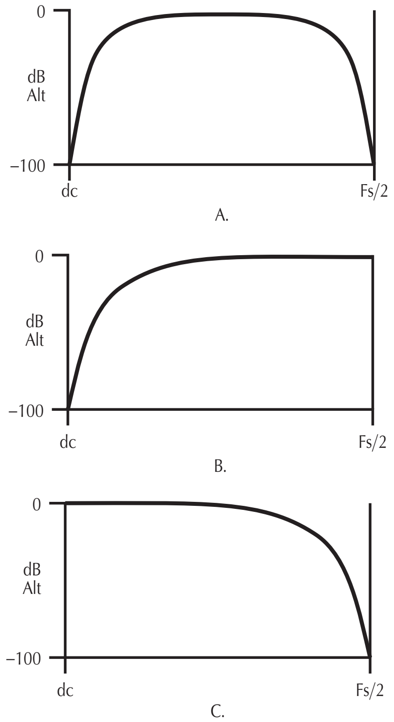


Figure 25-137. 3 point FIR filters from b0, b1, and b2 of Fig. 25-136.

seen clearly; the b0 and b2 coefficients are really quite small; the reciprocal of this value is the amount of gain being generated in the IIR output chain. Thankfully, commercial design packages and most cookbook coefficient calculation routines take this scaling into account. But the underlying issue is quite serious. Using, say, 0.0001 as a b0 coefficient (not unrealistic) and assuming a maximum input signal of 1 (the maximum signal range using the fractional arithmetic scheme in some fixed-point DSPs is 1 to –1), then a value of 0.0001 will end up in the accumulator, and despite the huge feedack-derived gain in the IIR output chain, the contribution to the output from the input signal is still only 0.0001; this corresponds to –80 dB. If the output were to be truncated to 24 bits (144 dB) the bottom 13 or so bits worth of the input signal would effectively, be sawn off and thrown away, leaving us with an 11-bit system. In numbers, this leaves a maximum signal to floor ratio of only 64 dB; if the normal operating level of the system is –20 dBFS (deciBels below Full Scale) (0.1), that is only –44 dB signal to floor. Practically, with rounding, noise shaping , or dithering, things are worse.

The good news is that such quantization noise in filters can sometimes be masked somewhat by the very signals the filter is passing; the bad news is that when it is audible, it is Audible. And even when not overtly audible, it lends itself to a disquieting roughness to the sound that is difficult to pin down. The accumulator has all the width needed for valid data; standard practice on any filter on which this is even likely to be an issue is to make the IIR chain delay storage wide enough to fully encompass the attenuated input signals. In the specific example of a 24 bit fixed-point processor, the IIR output delay chain is made long, or double width at 48 bits. It also means that nearly always the a1 and a 2 IIR multiplies need to be long too—i.e. the lower 24 bits need to be MAC'ed in with the upper 24, which increases execution time of the filter.

Sharc and other floater programmers are permitted to smile at this point. It is a happy day to doff the shackles of fixed point.

25.21.2.5 Cascaded Biquads

Having previously noted that IIR filters with more than a biquad's worth of delays and multiplies are not attractive, there are approaches to coupling more than one biquad with the intention of making more complex or effective filters or simply those of a higher order. Better than just running one after another. Fig. 25-138 shows such an arrangement; the second biquad uses the output delays of the first as its input delays, and so on.

25.21.3 Parametric EQ

Raw biquads can take care of most traditional filtering. One approach to doing a console-style parametric EQ section, with independent control over center frequency and *Q* of the employed filter and of the amount of lift or cut introduced, is shown in Fig. 25-139. A standard biquad is fed directly from the source audio, which is also attenuated by (in this case) 12 dB by the expedient of arithmetically shifting the data two bits to the right (down), or by multiplying by 0.25, in the DSP. The filter's output, fed through an attenuator, is summed with the attenuated direct signal, and the result arithmetically shifted (ASL) two bits to the left (up 12 dB). This shifting up and down allows a correspondingly higher amount of the filter to be present in the output, which is required if high levels of boost are required. This example's 12 dB allows a maximum boost of 13.8 dB to be achieved, which happily encompasses the ±12 dB control range often found in EQs; more boost capability would require greater shifting down and back up. One can in a floater (floating-point DSP) leave the straight signal alone and simply multiply the filter

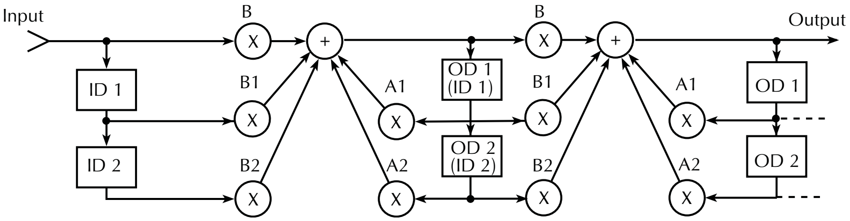


Figure 25-138. Cascaded DF 1 Biquads, sharing delay lines.

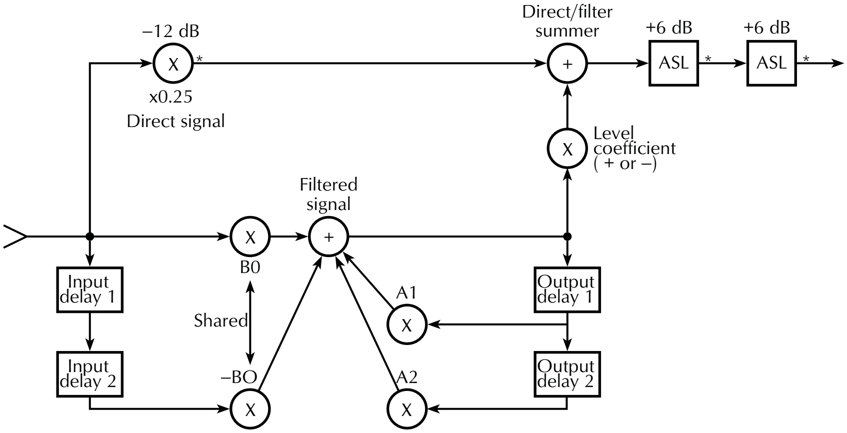


Figure 25-139. A bandpass parametric EQ stage.

output up as much as is necessary (the attenuator becomes a gain stage) and avoid the shifts entirely.

EQ boost is achieved by adding in filter; cut is done by subtracting it away—a negative coefficient is thrown at the post filter attenuator instead of a positive one. There is a non obvious criterion for cut coefficients—as one cuts, the effective *Q* of the EQ responder tends to sharpen; the frequency response of this arrangement at, say, 12 dB of cut is not complementary to that at 12 dB of boost; one needs to relate and modify the filter *Q* with cut level in order to take this into account and retain customary lift/cut symmetry.

Multiple sections of parametric EQ can be and usually are simply cascaded, although emulation of many classic analog designs has been better served by running the multiple filters in parallel and then adding their gained results all together with the straight signal. The band interactions are entirely different, offensive to a tidy mind, but far closer to the truth!

Given that most parametric EQs use bandpass filters only (at a push even shelving filters can be faked reasonably well using such) and that, as we've seen, bandpass filters have their b1 coefficient always at zero, it can make sense not to perform that multiply at all, thus saving data fetches and a multiply. Additionally, since the b0 and b2 coefficients are simply inverse of each other, only one need be sent from the host processor to the DSP, the inversion being simply achieved internally. This is welcome streamlining of the processing.

25.21.4 Shelving EQ

Real shelving can be achieved by using a full biquad in the EQ (as opposed to the simplified bandpass-only variety shown) with low-order high-pass or low-pass filter coefficient sets, or an even simpler structure as in Fig. 25-140. Much greater than a single-order response in the filter tends toward a frequency response with a "phase-bounce" in it near the turnover frequency, generally considered undesirable (except perhaps when one is being very picky emulating a Baxandall). The arrangement shown is a shelving EQ using very short filters. Advantage is taken of the fact that with single-order filters one can very easily create a high-pass filter merely by subtracting away a low-pass from a straight signal.

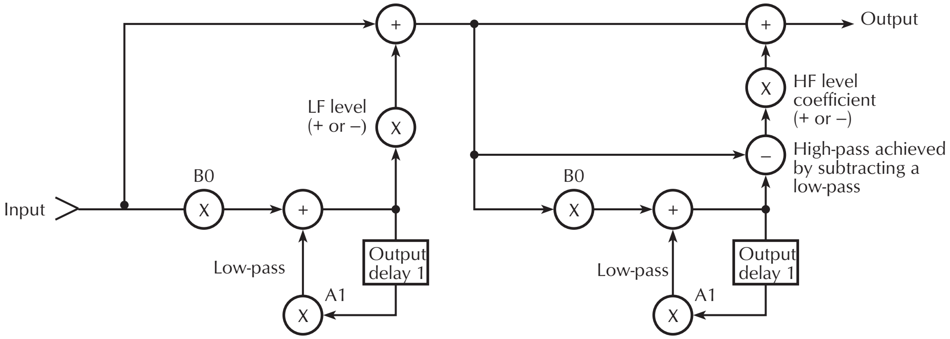
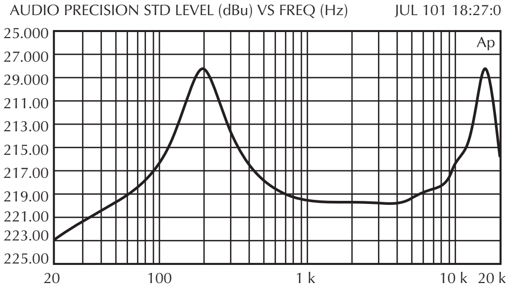


Figure 25-140. Shelving EQ using single-order filters.

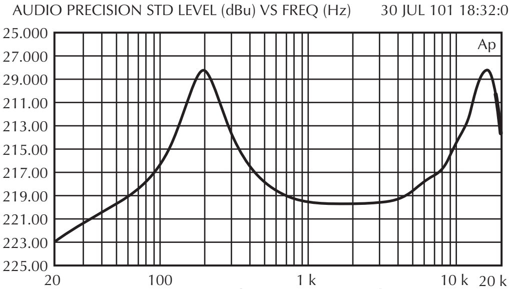
25.21.5 EQ High-Frequency Response Anomalies

Odd phenomena occur when filters are attempted too close to half the sample rate, 24 kHz in a 48 kHz system for discussion here: partly as a result of the inevitable zero in response at half the sample rate with bandpass filters, as we saw, and partly an effect of the prewarping in the transform calculations used to create the filter coefficients. The effective *Q* of a filter as used in a parametric such as this appears to increase (become sharper) and become asymmetric (high side gets steeper) as its curve approaches Fs/2. Although this effect can be considered unimportant, occurring at the audible extreme as it does, this behavior can be improved by the expedient of applying a subsidiary correction to the desired *Q* value prior to warping, or by the more fundamental approach of oversampling. This basically means running the EQ (or at least the HF bits of the EQ) at twice the sample rate; upsampling to 96 kHz and downsampling (to get back to 48 kHz) are quite straightforward. This has the effect of pushing the squiffy zone up toward the new Fs/2 of 48 kHz, where it simply won't matter, keeping the normal audio-frequency range of EQ linear and tame. Under some conditions with some program material, upsampled EQ (even though subsequently brought back down again) can sound better. One has to be very careful with the nature of the reconstruction filters in the upsampling in order not to imbue even worse funnies in EQ frequency response than one is trying to fix.

Fig. 25-141A shows the squiffy effect on a 16 kHz *Q* of 2 parametric EQ section; a similar *Q* of 2 filter at 200 Hz is shown for comparison. Correction (not oversampling in this case) results in the improved lower-frequency slope of the 16 kHz filter; this is now comparable to the skirts of the 200 Hz filter (Fig. 25-141B. Unfortunately, there's not a whole lot one can do about that zero at 24 kHz without oversampling, so EQ close up to the band limit will always be a bit suspect.



A. Uncorrected HF EQ anomaly.



B. Corrected HF EQ anomaly.

Figure 25-141. High-frequency anomaly

25.22 Digital Dynamics

There are many approaches to dynamics processing in digital, but most fall under one of two categories: mapping and literal. Briefly, *Mapping* involves creating a plot, a table, or a map describing what the desired output level for any particular input level needs to be; an input sample comes along and based on its value, a gain-control value is picked out of the look-up table map that is to be applied to that particular value of input in order to create the desired output level. The map can contain the transfer values for many different sorts of dynamics processing simultaneously—say, compression and gate, limiter and expander, etc. Fig. 25-75 in the earlier discussion of dynamics gives a clue as to the structure of such a map.

Literal is building a digital processing equivalent of how one would literally achieve the processing in good old analog.

For all the fuss that is made about how DSP makes doing audio design harder, it's nice to come across aspects that make life so much easier and nicer. Here are two:

- A machine language mnemonic for ABSolute value. The effect is to look at a number and if it's negative, make it positive. It's the DSP equivalent of a precision rectifier, never the most trivial of analog design exercises.
- MPY (Multiply)—This is the most perfect, distortion-free, vice-free gain-control element. One will never, ever, want to play with VCAs or FETs again.

25.22.1 Mapping Dynamics

Look-up table dynamics have the strong property of being very fast in terms of processing cycles at the expense of the memory for an adequate map or set of maps. The precalculation has already been done, so all that needs to occur is the indexing of the look-up table from the value of the input signal—returned is the gain-control value—very short and to the point. Depending on the dynamic range over which the dynamics is to behave, the depth to which (how low in level) the dynamic behavior is to be adequately described (important for gates, expanders) and importantly the degree of resolution of the table (so that signal levels don't noticeably lurch from one value to the next), the size of the tables can get quite healthy. In addition, it is often convenient to actually run more than one table. Required memory usage may or may not be a problem—some DSPs have huge rafts of memory, while others, designed for less memory-intensive streaming audio applications, may have only just enough for the basics.

The map only describes the instantaneous gain value. Direct application of recovered gain values would result in awful distortion. Obviously some temporal constraints need to be added. Typically these are the classic dynamics values of attack and release and such. Where these time constants are applied is an interesting question. In order to use the usually relatively slow release time constant to smooth out the inevitable steps from the table quantization, this usually follows the look-up. If one were to be emulating a peak limiter, then one might well let the input signal directly pick its value from the table and then apply the usual short attack time constant to that. In other words, for a limiter, both attack and release would follow the look-up stage.

Compressors generally have a far more relaxed attack time, with the intention of deriving a signal more corresponding to the audio energy than its instantaneous peak. In this case the attack processing would take the form of short(ish)-term averaging or even rms-like detection; the result of this averaging would be used as the pointer into the look-up table. The release would be left on the output of the look-up, mostly for its role as janitor, tidying up the potentially ragged steps.

Assuming a compressor, a likely threshold range would be from −40 dB below nominal operating level up to, say, 10 dB above. Since nominal operating level is usually at or around −20 dBFS, this implies that the look-up table has to encompass an input signal range of −10 dBFS down to −60 dBFS. It has to do this with sufficient resolution that no gain lurches are obvious. (Although most musical program material can withstand even comically large gain lurches under these circumstances, some—solo flute or a slowly decaying tremolo bass-guitar note spring to mind—will highlight painfully small ones.) Since the gain steps should almost certainly be dB linear or close to it, and the applied signal is linear, it is wise to perform a logarithmic conversion to the input signal to closer approach dB-to-dB mapping in the table. These tend to be computationally expensive or iinvolve look-up tables themselves (!), but the penalties for not prelogging are either a look-up table to achieve adequate resolution at the lower levels (and −60 dBFS is a long way down, to 1/1000 in fact) or reduced accuracy at the lower levels for a smaller map. A linear map for this compressor might need to be 2048 or 4096 steps deep to have nonembarrassing behavior near the bottom.

Big tables, actually big anythings, are bad news in the sense that if the parameters are changed (say the compression ratio is altered a notch) a whole whacking great new table has to be fed from the host microcontroller up to the DSP. The alternative, if the memory in the processor supports it, is to permanently have a suite of maps encompassing the range of parameters required. A different map is pointed to when a parameter changes. LOTS of memory!

A nice lateral thought solution to the really deep map problem affords itself with the use of floater (floating-point) processors. (Actually, it can be and is applied in literal approaches, too.) This is to create a good and concise map, or set of maps for various changes in parameters, and then to move the threshold around by scaling up and down the actual audio samples accordingly. In particular, expansion curves can be created economically of memory; instead of moving the curves around, the audio is scaled instead to create the desired response.

A peak limiter, operating over a comparatively much reduced dynamic range, may possibly eschew a log convertor and just look up directly. On the other hand, an expander may have to adequately describe down to −90 dBFS (or whatever the "don't care" level might be). Which brings to the fore another point, which is that if different time constants are required for different functions, as they certainly would be between a compressor and a gate, say, it could make sense to use a different look-up table for each.

25.22.2 Literal Dynamics

This is the technique of emulating (as close as one can) how an analog circuit achieves the required dynamic behavior. There is a bit more art in this approach, and although the algorithms tend to be longer and certainly more intensive than mapping, there is very little memory usage, and changing parameters just involves sending a handful of coefficients to the DSP from the host, rather than potentially thousands.

It is possible to emulate the rough-and-tumble free-for-all uncontrolled servo-loop behavior of a feedback-style compressor/limiter, or alternatively plod through the tidy-mind deterministic feed-forward VCA approach, which involves division and/or much loging, antilog'ing, and untold processor time (transcendental functions are very long-winded in DSP), for ultimately a well-behaved but, frankly, bland result. (Guess which the author finds more fun?) Filling a whole DSP with such a VCA-like processor isn't difficult.

There is just as much latitude for approach with literal dynamics as there necessarily has been with analog design; indeed, if one's goal is to emulate classic analog dynamics this is really the only way to go.

25.22.2.1 A Simple Digital Limiter

Fig. 25-142 highlights how dynamics signal processing in DSP—in this case a simple peak limiter—can almost slavishly follow an analog architecture.

The key to the limiter's operation is the gain-reduction value—sorry, the author still thinks of this as a control voltage. Remember that multiplying a signalcoefficients by 1 doesn't change the signal; multiplying by a fraction less than 1 reduces the output signal—i.e. affects gain reduction, which is what we need when a limiter is biting.

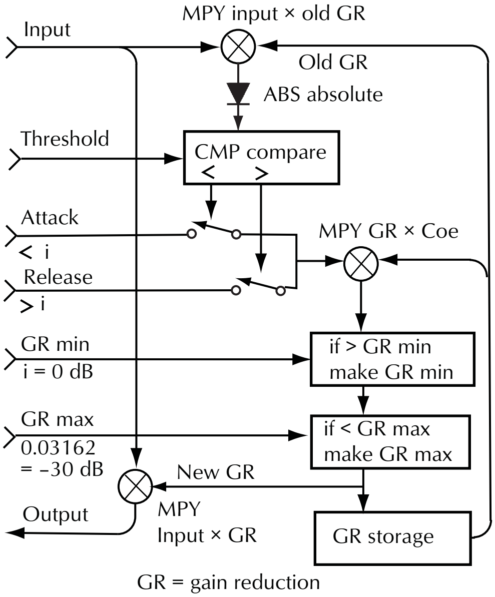


Figure 25-142. Flow chart of a simple digital limiter.

First, the immediate present (new) sample is multiplied (MPY) by the stored GR-generated last sample. This is necessary to judge whether and which way this last GR value needs to be adjusted for the present sample. The absolute value (ABS) of this modified input sample is then compared (CMP) to the threshold coefficient. If it is greater than the threshold the GR value needs to be reduced, and the program branches into attack, where the old GR value is multiplied by a coefficient usually just slightly less than 1. Likewise, if the threshold isn't breached the GR value can be relaxed, so it branches off to release, where it is in effect multiplied by a coefficient just ever so slightly greater than 1. This is shown in the diagram as switching between using the attack coefficients, or the release coefficients. Naturally, the modified GR value has to be clamped such that it can't rise higher than 1 (and so no longer be GR!) and that is also the normal unity "resting" case.

The coefficients for attack and release are in this simplistic case (it can get considerably more complex!) multiplicative—the GR value is changed by the same proportion or, in other words, the same number of fractional dB per sample. Running at, say, a 48 k sample rate, in order to have a 1 dB/s release rate the coefficient would have to represent 1/48,000 dB increase in GR value, or about 1.00005. "Slightly" says it all. This dB-per-time gain trajectory works quite well in audio, emulating the dynamic response of many good analog systems.

The last things that happen are that the newly modified GR value is saved for use next sample and also used to multiply with the present input sample to create the gain-reduced output sample.

All in all, it is almost an exact parallel to a simple analog feedback-style limiter; complexity concessions exist for operating in a sampled-time system (such as the initial input/last GR premultiply), as opposed to relying on the always existent signals in the continuum of analog. On the other hand, one effortlessly achieves true dB/time gain rates for attack and release, usually a feature of posher analog designs and only ever approximated in simple systems.

25.22.2.2 Feedback-Style Limiting and Compression

Unlike most analog GR elements, the "MPY" in a DSP is directly linear in operation—i.e., a gain-reduction value of say 0.3 will cause the signal through the multiplier to be reduced some 10 dB. It is not linear-by-dB, like a VCA, or a mangled exponential/logarithmic like thing such as from using a raw semiconductor element such as a transistor or FET. Yet, as has been shown above, linear-by-dB results can be achieved fairly simply. Emulating other laws can get rather interesting, but are certainly attainable, in pursuit of a sound. Similarly, the determination of the amount of instantaneous feedback in a feedback limiter depends on many things, not least the attack and release time constants necessarily applied to it. Another is whether the control signal is being generated all the time and only applied when the threshold is exceeded or alternatively if the control-signal determination is only woken up when the threshold is exceeded. Both can work well, and both sound utterly different.

Feedback-style compressors can use basic limiters as above as a starting point. The limiter (using the required attack characteristics of the compressor as its attack/release time constants) creates an overage signal implicit in its own control signal, representing the amount the input signal is exceeding the limiter threshold at a given moment. By manipulating this overage so as to create a control signal more in accord with a chosen compression ratio rather than the hard limiting, a suitable release time constant applied, the doctored control signal is used in a second multiply on the untrammeled input signal, outside of the feedback loop. As an approach, this combines the edge and sound of a feedback-style dynamics unit with a sane deterministic compression ratio. Fig. 25-143 shows a family of deceptively analog-looking input output curves from a digital soft-knee compressor using the described technique.

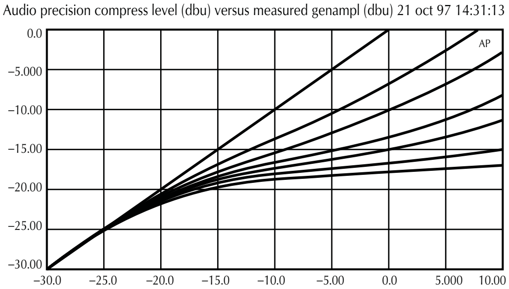


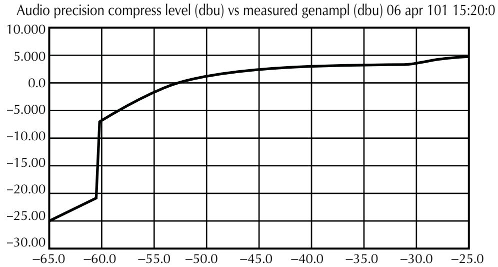
Figure 25-143. A family of compression curves from a digital dynamics section.

25.22.2.3 Gating

The purpose of a gate is to attenuate completely or partly a signal that falls below a given threshold. Typically they should wake up (open) quickly, hang open for a while if the signal goes away just in case it really hasn't, and then close at a gentler rate. Also, to prevent "falsing," there are two thresholds, one for opening and the other, slightly lower in level to determine closure. Written as described above, it is about as "digital," yes/no, a set of conditions as one can ever hope to meet in audio and is a complete natural for the literal approach.

The absolute value of the input signal is compared to the open threshold; if tripped, a target control signal of 1 (unattenuated) is applied to a short low-pass filter bearing the attack (open) time constants feeding the attenuator multiplier (this will quickly ramp up the gate to open); at the same time a counter is initialized. The counter is the hang-time counter. It is reinitialized at every sample that the close threshold is exceeded such that it doesn't get a chance to start counting down unless the signal really has gone away. If that occurs, and the counter does count down to zero, a control signal value of zero (for off) or some other value representing an amount of off attenuation (depth) is applied to a longer release time constant low-pass filter, the output of which is applied to the attenuator multiplier.

Fig. 25-144 shows the dynamic transfer characteristics of a microphone input using a combination gate/soft limiter; this combination is used extensively, in this case for a stage backup vocalist. If the singer is making enough noise, the gate opens (actually it relieves 14 dB of attenuation, which is enough to make stage spill go away enough), and the limiter almost immediately takes over, keeping the voice at a manageable, consistent level for the mix. Note the 3 dB of makeup gain. Lest one is concerned that this combination is far too unsubtle for quieter songs, remember that this is a digital process and in the context of a programmable console can be (and is) reprogrammed to suit on a song-by-song or even section-by-section within a song.



Soft-knee limiter and gate characteristics I/P gain at 50 dB; gate opens at −60 (−30 dBFS), limiter nominal knee at −50 (−20 dBFS), clips at −30 (0 dBFS).

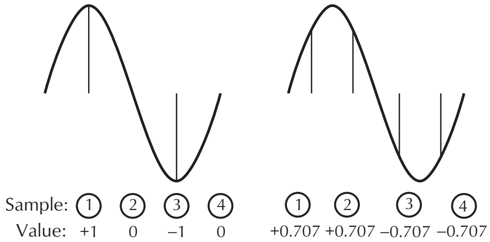
Figure 25-144. Soft-knee limiter and gate characteristics.

25.22.3 Inadequate Samples for Limiters

By and large, Mr. Shannon and Mr. Nyquist did great jobs. Digital audio works really well and it is indeed possible to reconstruct an indistinguishable result from a source through a system sampling barely twice as high as audio bandwidth. But there are a couple of places where a limited number of samples trips up—In particular, attempting to sense the peak of a signal, as one attempts to do with limiters and gates, both of which need to respond accurately to them.

Unfairly and unreasonably (since it isn't terribly relevant to most real audio) we'll consider first the example of Fs/2—half the sample rate. With only two samples per cycle of sampled signal it is entirely possible for the two samples to miss the signal altogether, if they happen to occur at zero-crossing points of the applied audio signal. But, then again, they might hit the jackpot at the crests.

More realistically, look at the two extreme cases of Fs/4, or 12 kHz for a 48 kHz sample rate, in Fig. 25-145. Nobody is arguing that 12 kHz isn't audible, yet here is a case where there can be as much as 3 dB error in sensing a level. There are similar, if far less serious points of error dotted throughout the audio spectrum (e.g., Fs/8, 6 kHz etc., or anywhere else where Fs is divided by an even number). Now, to be completely fair, under any reasonable circumstance this effect would not be excited and certainly not be audible, since an exactly 12 kHz tone in isolation is not by and large terribly common or useful—And would be a far-reached argument in support of a blanket increase in sampling rates for digital audio. However, in the specific case of attempting to peak-sense audio levels, one comes across these spot frequencies with too little effort. This is a reconstruction error, or more precisely, an error due to the samples not explicitly describing the signal but relying on a later reconstruction filter to fill in the gaps. Back to analog, the signal reconstructs just fine! This is just a hint of the occasional disconnect between the two domains.



A. Lucky. B. Unlucky.

Figure 25-145. Sampling Irregularity at Fs/4 (12 kHz); possible 3 dB error.

A second—and actually practically more worrisome—reconstruction error effect is if through clipping or heavy dynamics processing a pair of adjacent-in-time samples are made full-scale, a downstream reconstruction filter will cause a significant overshoot beyond full scale in the recovered analog signal. This can cause clipping in the following analog stages if insufficient head room is allowed, and a nasty surprise to anyone who thinks full scale is the most one can see out of a digital system!

By way of a slightly different practical example of sampling oddities, it was noticed that digital limiters seemed to respond differently each time to a snare-drum impulse; sometimes they'd catch it hard, sometimes they wouldn't. In comparison an analog limiter just, well, caught it. Now, snare drum is pretty evil, a nasty big initial short spike. On analysis, sometimes the spike was adequately described by the limited number of samples, sometimes it blew through the gaps. There wasn't much difference, between 3 and 6 dB in captured peak level, but that is plenty enough a difference to be audible and, more to the point here, plenty enough to invalidate the devices as peak limiters!

Oversampling—i.e. making the sample rate twice or even more times higher—has the effect of pushing the worst of the reconstruction-error potholes out of the relevant audio band. Even though one is operating on exactly the same audio data originally sampled at the lower inadequate sample rate, peaks are captured accurately enough for all practical purposes. The missing peaks of Fig. 25-145 are actually being filled in by the reconstruction effect of the low-pass filter employed by the upsampler, in exactly the same way as a D/A's reconstruction filter would.

As such, there is a very strong argument for oversampling in peak-limiter and other fast dynamics processing. Since it has one tightly defined purpose, the impact of doing the sample-rate conversions does not have too great an impact on DSP processing cycle budgets.

Unfortunately, for the live circumstance of that example snare drum, actually initially converting at the higher sample rate is unavoidable if a more analog behavior is required; the greater number of samples affords fewer gaps for those nasty transients to escape through.

25.22.4 PreDelays, or Look-Ahead

Applied only in rare cases in analog because of the difficulties in providing for good audio delay, predelaying is eminently achievable in digital dynamics sections. Predelay is the technique where the main signal path through the dynamics is delayed for a short period (1 ms, 2 ms or so) to allow the side-chain processing to determine the right amount of gain reduction to be applied; this value is then applied to the main signal path in a gain-control element discrete from that used in the sidechain, Fig. 25-146. The prime use is in peak limiters (which are nearly always feedback style, even where the other sections may be feed forward), where overshoot, which can occur during this onset settling period, can be completely avoided. An improvement in sound results, too, since very hard and brutally short attack times can be mellowed out knowing that overshoot is not going to increase as a result. A relatively soft attack time (for a peak limiter) of 1 ms combined with a comparable predelay captures the peaks without the need for subsidiary clipping and yet is sufficiently aggressive that it retains its loud characteristic but without the usual telltale ripping hard edge."

Look-ahead limiting is extensively used in broadcast air-chain processing, and especially on feeds to streaming compression codecs (AAC, MP3, or HD radio, for example), which generally do not react well to the artifacts generated by more conventional clipping or unavoidable transient escapee overloads from ordinary limiters.

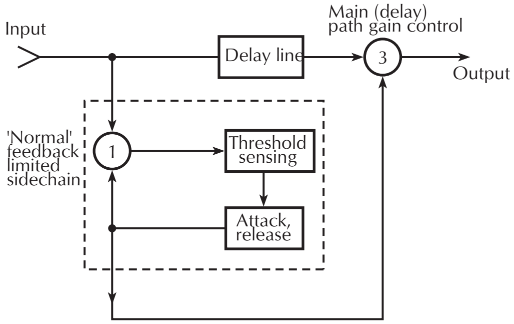


Figure 25-146. Dynamics section pre-delay system.

Only occasionally would such processing be done in a console channel, but when it is, it should be remembered to apply equal delay (whether limiting or not) to other contributory channels in the mix.

25.23 Digital Mixer Architectures

Two distinct approaches seem to be taken to the signal-processing architecture in mixing consoles, probably stemming from how deeply steeped in traditional computer science the designer is.

25.23.1 The Sea of DSPs approach (sometimes known as a Sharc Tank).

In this, a large enough array of DSPs for all envisaged processing is closely coupled to enable the rapid transfer and sharing of data between them. The "tank" is fed with all sources, and all destinations are taken from it. In a telephone-exchange kind of approach, signals requiring processing are farmed out to other processors in the tank and the results returned; it could be regarded from the outside as one big processor. The main advantages of this approach are that not being physically constrained to a particular organization, reconfiguration is straightforward; any signal can go anywhere at any time for any purpose. If more processing is needed it is merely attached to the busing system, growing as required. The major downside is that all the flexibility makes programming such a beast very difficult (the word *nightmare* has been bandied around).

The "one big processor," is of course a reality for many modest-sized applications, in the form of the humble PC. These have increasingly faster and more capable processing cores, and multiples of those, too. Although they have their own issues as far as processing audio (e.g., they generally don't do DSP terribly efficiently), brute force, speed, and might make up for those. A major positive is having everything under one roof, dispelling the problems of multiple-part interconnectivity.

The second approach very much follows the signal-flow approach of a conventional analog mixing console, using multiple DSPs as required, processing being applied in-line as and where it is needed. On very large consoles the mixing processing itself can take on the look and feel of a tank-style array, but other than that the layout of the signal paths has remarkable parallels to analog.

25.23.2 A Practical Digital Mixer

As with the discussion of analog consoles, which revolved around the description of a particular design, so this section uses as its basis the architecture of a real digital mixing console. It is shown in its basic form. As the reader is well aware, pin-by-pin details of implementation, bells, whistles etc. can rapidly mushroom and the weight of resulting detail tend to obscure; as it's not too difficult to figure out how most of this is done, they have been omitted for clarity. It is important to remember that in terms of lines carrying audio signals, it is accurate, due to the use of the serial audio format outlined below. Shown here in a mid-size 64-by-24 format, this particular design's premises were simplicity and scalability (it can be readily made bigger or smaller) and has proven to be robust and reliable, using no scary technology and with nothing running on the edge. Over the years this basic architecture has grown and evolved through generations of increasingly powerful DSP and support devices with the odd effect in this blighted world that it has actually become progressively simpler to build with time. Also the steady and welcome improvement in integrated converters has resulted in the overall performance blossoming to the extent that this, along with other digital mixer designs using comparable technology, owe nothing to analog in performance whatsoever.

It is assumed, of course, that the control surface has been undertaken as the separate design exercise that it largely is; this discussion concerns the signal-processing side of things.

25.23.2.1 Serial Audio Format

Nearly all converters and like peripherals such as AES/EBU format transmitters and receivers use in common a serial digital interface; this is usually set up as to be two sets (left and right) of 32 data bits per sample frame (64 total), meaning a data rate of 3.072 MHz (for a 48 kHz sample rate). This is a very tame and robust rate and can be run around quite happily without fear of corruption, and as such is used as the nearly sole means moving audio data around in this console. Adopting this serial format also minimizes the amount of data format changes required.

25.23.2.2 Inputs

Input signals are applied to whatever form of convertor or interface is required: microphone amplifiers or line-level inputs into A/D converters, AES/EBU into AES receivers, and subsequent sample rate converters. Sample-rate converters (SRCs) are necessary since it is unlikely (unless a whole amount of trouble has been gone through to synchronize the whole system of which the console is a part) that other digital sources will be and remain in word/data-rate synchronization with the console. The recorder may well be, but typical AES/EBU devices, such as outboard effects, or remote sources, rarely will be. If it is considered necessary (on the basis that anything that messes with data unnecessarily is a bad thing), the SRCs may be bypassed for sync'ed sources, but frankly SRC's today have artifact levels so low as to be considered quite blameless.

At this point, all the data is in native format (the convertor serial standard), travelling in pairs—mono signals (microphones, say) in pairs and stereo sources as left/right pairs per data line. For a 64 input console, this means 32 data lines.

The channel signal processing is done four channels (two pairs) per input DSP, Fig. 25-147. The DSPs used here very conveniently have native format inputs and outputs (being designed to work with normal converters), making interfacing really simple. They are also easily powerful enough to do four well-featured channels worth of signal processing. Typically, this would be high-and low-pass filters, a four-band parametric EQ and limiter/compressor/gate dynamics, and delay (memory is attached to the external memory interface of the device to support this if required). The channel DSP has spare input and output capability, which can be implemented if required as selectable direct channel outputs, keying inputs to dynamics, etc.

25.23.2.3 Mix Stage

The 64 channel outputs are taken from the 16 channel processing DSPs as 32 output lines and applied to the mixer stage(s). Fig. 25-148.

The ominously large device labeled FPGA (field-programmable gate array) into which all those lines

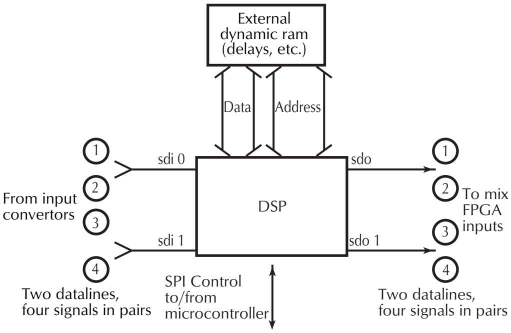

External
dynamic ram
(delays, etc.)
Data
Address
①
②
③
④
From input
convertors
sdi 0
sdi 1
DSP
sdo
sdo 1
①
②
③
④
To mix
FPGA
inputs
Two datalines,
four signals in pairs
SPI Control
to/from
microcontroller
Two datalines,
four signals in pairs

Figure 25-147. An input DSP (one of 16).

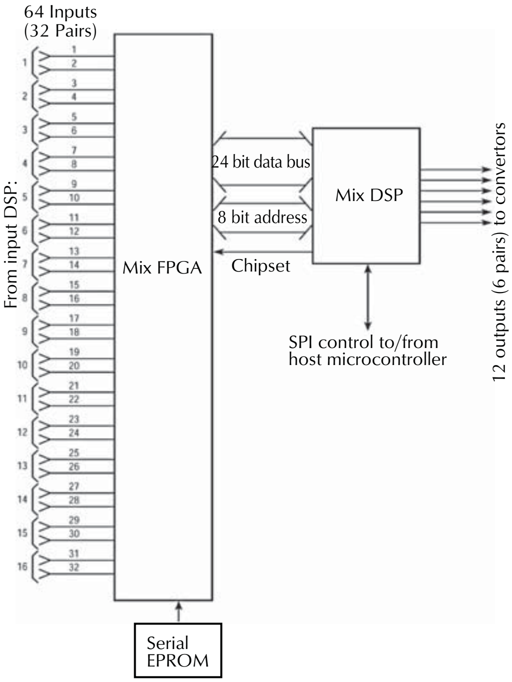

64 Inputs
(32 Pairs)
From input DSP:
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
Mix FPGA
24 bit data bus
8 bit address
Chipset
Mix DSP
SPI control to/from
host microcontroller
12 outputs (6 pairs) to convertors
Serial
EPROM

Figure 25-148. Mix stage.

disappear is programmed to be merely a (large) collection of serial-to-parallel data converters—no voodoo. A slightly simplified version of it's contents is shown in Fig. 25-149. The FPGA takes each data line and puts it into its own 24 bit long shift-register; when it has counted that the necessary 24 bits have arrived, it seizes the data and tells the mix DSP with which it is associated that the data is ready for harvesting. (A long-ago prototype of this design actually used discrete logic shift registers. Lots of them. It was huge. FPGAs are much better.) To the DSP, the 32 shift register outputs are arrayed and addressed to look exactly like memory, and indeed, the FPGA sits on the 24 bit wide external memory bus of the DSP, with enough address lines to uniquely address each shift register location. Once informed the data is ready, the DSP copies the data values down into its own internal memory, from which the mix code accesses it. Although this can be done in real DSP software, it is usual to invoke a DMA routine (direct memory access) that, depending on the sophistication of the chip, can transfer data quietly in the background of normal processing from one area or peripheral into/out of internal memory with minimal impact on normal operation. In practice, it always seems to slow things up a bit (*background* is a relative term, it seems), but overall, DMA is slicker. The FPGA/DSP DMA combination does this transfer operation twice per sample, once for left data, the other for right. These two sets of data are held in buffers in DSP memory so as to be in time alignment ready for the next pass of the mix code.

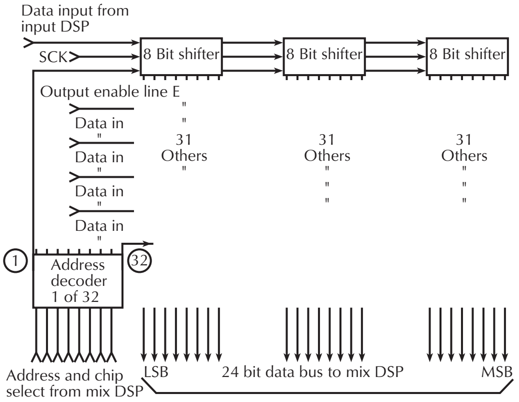

Data input from
input DSP
SCK
8 Bit shifter
8 Bit shifter
8 Bit shifter
Output enable line E
Data in
Data in
Data in
Data in
31
Others
31
Others
31
Others
①
㉜
Address
decoder
1 of 32
Address and chip
select from mix DSP
LSB
24 bit data bus to mix DSP
MSB

Figure 25-149. Simplified contents of FPGA.

25.23.2.4 Mix Code

As earlier mentioned, DSPs are designed to do some functions really well and one of those is the FIR filter. This involves multiplying a piece of data by a unique coefficient, adding the product into an accumulator and then rapidly moving on to do the whole thing all over again (next data point, next coefficient), and again, for as long as the filter may be. Well, from a mixing point of view, a group output multiplies an input channel sample by a unique coefficient, adds the product into an accumulator, and then rapidly moves on to do the whole thing all over again (next input sample, next coefficient), and again, until it's done all the input channel samples. Got it? A mixer and an FIR are as far as the

DSP is concerned chemically indistinguishable. For each mix group in turn the DSP addressing runs through all the input data in turn, multiplying each by its appropriate gain coefficient. Automated addressing in the DSP keeps track of which coefficient goes with what sample. It is about as efficient a processing operation as it can possibly be.

There are of course limits as to how many channels and how many groups a single DSP can mix:

Crosspoints. Each channel-to-group calculation is a crosspoint (from the concept of the mixer being a big soft matrix). A 150 MHz DSP has a little over 3000 processing cycles at 48 kHz, but necessary programming overhead precludes the use of all of these for mixing. Around 2000 crosspoints is perhaps more realistic. So for 64 inputs, 32 output mixes per DSP is theoretically possible; 32 sources, 64 outputs; 128 sources, 16 outputs, and so forth. This elasticity has limits, however, as follows:

Input Bounding. This is the limit on how many sets of input data one can realistically capture and shovel down into the DSP and still leave it with processing time to do any mixing. Even DMAs have some time impact—in general, DMA notwithstanding, the more time taken dragging data around, the less time available for processing. External memory fetches take time (access to such memory is much slower than to internal memory—hence the need to bring the data down from the FPGA rather than access it directly for the mix). Even though it is theoretically possible to spend an entire sample period wheeling in new data in the background for calculation in the next, data management can get pretty hairy. In comparison the actual mixing is a doddle.

Output Bounding. This is less of a problem, since by and large there are fewer of them. But expecting a large number can lead to an issue of how to get all those mixes back out into the world.

Since simplicity was a major aim of this particular design, the outputs of the mixer stage are taken as six pairs (twelve groups) from the DSPs in-built serial interface; these group outputs are applied in the same mix'n'match fashion as the inputs, directly to whatever class of output device is necessary—D/As for analog outputs, AES transmitters for digital, etc. Deriving more outputs from the DSP involves getting those mixes back up into the FPGA—again by DMA—and doing a parallel-to-serial conversion of each there, in reverse fashion to that done on the inputs to the mixer stage. By such means, the modest processing power in this mixer core can easily handle a 64-by-32 console. Here, the twelve buses not coming directly serially out of the DSP are done in this manner.

Increasing the number of mix buses yet further would be achieved by using another FPGA/DSP core, allowing a further thirty two mix buses. The second mix stage's FPGA is simply parallel-fed from exactly the same thirty two data lines from the input channel processing DSPs as the first mix stage.

The fact that the mix outputs are in serial native format dramatically facilitates the dropping in of a further DSP for post mix processing if required for some applications–graphic EQs and group dynamics sections, for example.

As can be seen from the diagramS and the description, the signal flow of this console is in such striking accord with an analog implementation, one can rightly wonder what all the fuss is about.

FPGA's are becoming increasingly faster, more capacious, and capable with the addition of on-chip RAM and dedicated multipliers and such. A mixer of modest proportions, such as this design, is implementable directly into an FPGA alone with no need for the DSP. Currently, it is a cost-benefit design exercise, deciding whether a capable enough (and more expensive) FPGA is worth it over the low-cost DSP and cheaper FPGA. But the trend is clear.

25.23.2.5 Universal Mix Buses

The described design provides a large number of raw mixes, with no mix-specific hardware or code. It may be noticed that apparently an otherwise vital subsystem to the mixer appears to be missing—monitoring. Well, actually it isn't, and the fact that it is implicit in the design as it stands points out an approach and attitude to mix buses that would be hard to maintain in analog where every bus is a significant expense: in digital, buses come cheap.

Monitoring in this case commandeers a pair of mix buses (assuming stereo); think PFL bus for now. Any input to the mixer can be monitored on this bus by applying an on coefficient to the appropriate crosspoints to bring the source(s) onto the bus. So far so good. But for monitoring output buses (stereo group, auxes, any of them actually) rather than apply the analog solution of a selector to switch between those existing groups, what one can do is exactly recreate the mix to which one wishes to listen; if one were to apply the same coefficient set that is making, say, auxiliary bus 5 to the monitoring bus, too, then one will exactly recreate what is happening on aux bus 5 in the monitoring.

Where this approach really shines is main mix bus monitoring—one can mess about with the monitoring bus as much as one likes without affecting the main mix at all—nondestructive soloing becomes a reality, simply implemented at that.

Talkback is simply treated as one of the sources to the mixer; it can get routed into any of the mix outputs with no necessity of creating a separate subsystem, with IFB (Interruptible Foldback) talkover ducking or muting deriving from modified coefficients.

Note there is no distinction made between the group outputs as to what their ultimate purpose will be, group, aux, cleanfeed, etc. All that distinction is done at the control surface and the interpretation of its requirements by the host microcontroller—in other words the differences are all in the controlling software and not in the hardware that implements the mixes.

25.23.2.6 Coefficient Compounding

This is a rather fearsome title for a rather nice concept. This is how master fadering and group fadering are achieved. Rather than have a separate downstream gain stage after a mix has been achieved to effect overall level control, a convenient approach with a soft matrix mixer such as has been described here is to take the sensed level of the real, physical, group fader and then multiply each of the coefficients feeding that particular bus by its value. This is a direct analogy of VCA grouping, where one fader actually modifies the level contribution of each source to the mix bus, rather than gain changing the mix after the fact. Since all of these numbers (source contribution coefficients and group fader) exist in the host microcontroller, the arithmetic manipulation is quite straightforward. The database management aspect of this on a large console can get quite interesting, but this pseudo-VCA grouping approach is widespread and very powerful.

25.23.2.7 Coefficient Slewing

Rapidly altering coefficient data in a DSP runs into exactly the same tone click problem as do MDACs in analog; even small transitions made when the audio data sample is nonzero stand a very good chance of being heard as a click. A fader swipe can generate the all-famous zipper noise, and just as with MDACs, without care and attention the effect in EQs is little short of comical.

Sensing zero-crosses in digital is practically impossible, since particularly at high levels of high frequencies there may well not be any samples anywhere near zero—remember this is not a continuum like analog, the samples are just a regular set of stabs in the dark. A wide enough window to capture enough zero-crossings would probably be wide enough to still allow some transitions to be audible. Never mind the fact that the processing overhead for doing a window compare and decision on each and every coefficient would be overwhelming; it would probably cut the potential number of crosspoints in a mix stage down by an order.

A good solution is to allow the DSP to ramp relatively slowly between its present value and the new desired value, creating its own interpolating steps on a sample-by-sample basis small enough that each is inaudible. (This, by the way, is one of the necessary processing elements that eats up a chunk of mix-DSP cycles, limiting the maximum number of crosspoints available to significantly less than the raw cycles availability of the device would suggest.) A slightly different approach is to "pre-slew" the coefficients in an intermediate processor (often also a DSP) to offload the effort from both the host and the target DSPs. The inter-DSP communications can start to get a bit fierce, however.

It is a nerve-wracking moment when first trying on-DSP slewing. After all, the coefficients for IIR filters such as in EQs can be very, very touchy and have little tolerance for error before doing very odd things most unlike the filters they were intended to be. Amazingly though, it seems as though provided the filter set is stable where it starts, and stable where it ends up, it stays stable in between as the coefficients are slewed; it might get just a little wonky, but not enough to cause any serious sonic issues and certainly not enough to explode into what has been charmingly called "screeching cats from hell" (DSP audio guys and gals hear lots of them).

25.23.2.8 Clocking

A major subsystem within a digital mixer is clocking—making sure that each of the various circuit elements get the necessary hard, clean clocks required to operate properly. In this design alone there are six clocks for processing: 12.288 MHz master clock (actually divided down from 24.596 MHz to ensure symmetry), 6.144 MHz used as a master clock by AES/EBU transmitters, 3.072 MHz as the main serial bit clock for the standardized native serial data format, an inverse of that used by some A/D or D/A converters of less serial format flexibility than others, then 48 kHz, which of course is the data sample rate and houses left/right clock.

Although there is from a component-count standpoint a tendency to want to include the clocking generation in with an existing FPGA, say one from a mix stage, it can be beneficial to have it stand alone in a smaller FPGA or CPLD package. Generally, each clock feed to each device should be individually buffered and be as close to its target as possible. Needless to say, this takes a lot of FPGA/CPLD pins, and a single-purpose device starts looking like a good idea. The major benefit is that one can physically locate it where it can do the most good; this is as close as one can get it to the A/D, D/A and sample-rate converters. Ideally (but rarely is it possible) these should all be clustered in a "convertor ghetto" to keep the clock lines really short and tight from the clock generator, which minimizes noise and slewing on the various clocks, which can directly affect convertor jitter noise performance.

25.23.3 Signal-Processing Control

Fig. 25-150 outlines a typical control architecture for signal processing, or the *processing end*. It should be considered along with Fig. 25-124, which shows the *control-surface end*. The separation reflects that often the processing and the control are, indeed, in separate places interconnected by a network.

25.23.3.1 Controlling the DSPs

Each of the DSPs has an SPI (serial peripheral interface) port, an industry-standard means of device intercommunication. This consists on each device of a serial clock line, which synchronously clocks data in or out, and a serial data in line; these may be paralleled around all the DSPs. A serial data outline needs to be selected in a multiplexer for feeding data (such as metering information) back into the host processor. There is also a chip select line, which needs to be run individually back to the host; when yanked, a particular DSP knows that the data being clocked out on the serial data line is for it.

It is down this SPI interface bus that the DSPs receive their boot code at turn-on (the program code which it will run), a set of working coefficients (usually those that were current when the console was last turned off), and any changes to those coefficients as the console is being operated and parameters changed.

25.23.3.2 Metering

The indication to the user of the various channels' and groups' signal levels, dynamics gain reduction values, etc. is performed by the control-surface host, driving the appropriate indicators. How the data gets to that micro from the DSPs that are doing all the work can vary widely in implementation depending mostly on the physical configuration of the console. If it is a single box, with the signal processing under the hood of the control surface, then metering data can best be taken simply and directly from GPIO (general-purpose input and output) pins on the actual DSPs. Alternatively, but it is giving up a major advantage of the one box, the host micro could recover all the metering data from the DSPs and distribute it all accordingly. If, though, the

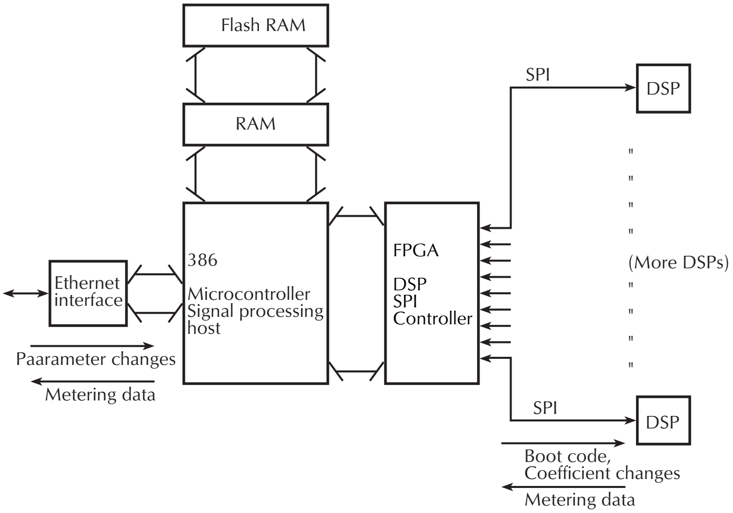


Figure 25-150. Signal-processing control architecture.

console is split, then a means of harvesting all the metering data from the DSPs, squirting it all up to the control surface, then disseminating it appropriately definitely has to be devised. This one sentence describes something that has many, many times been hopelessly underestimated, and at least in one case required a whole separate Ethernet run back to the control surface purely to handle metering.

In this design's case, the host micro polls each of the DSPs in turn, clocking back the metering information from each through the return path of its SPI; a packet is created of each complete console wide set, which is then delivered back to the control surface.

The good news is that metering data is not needed at anything like audio data sample rates. The feeds have been prefiltered in the DSPs with appropriate time constants and updating the relatively small data (8 bits is plenty) relatively slowly (better than every 25 ms or so) is adequate. Nevertheless, unless the polling by the host is under rigorous and deterministic control and the total bandwidth of even this fairly slow, small data set is carefully considered, the metering can start to be a major burden.

25.23.3.3 Host Microcontroller

This is usually a fairly fast and meaty micro, often of the x86 persuasion or large 68000 family. In the case of a self-contained console (control surface and signal processing all being in the same box) this will in all likelihood administer the control surface and displays in addition to the relevant function here, which is riding herd on the DSPs.

The host's job is to turn control parameters (as generated by the control surface) into coefficient sets that the DSPs can understand to perform the effect of those parameter changes. This would be by the coe-gen, or coefficient generation, software (typically written in C) and is equal in importance—a fact little appreciated—to the actual DSP code the DSPs are running. It is the coe-gen code that just as much determines how a console runs, feels, and sounds—after all, the DSPs are just doing what they're told and running code sent to them, by this host. By ways of example, the coe-gen code looks up what mix DSP crosspoints need to be modified to what coefficient values in response to a given fader being moved to a certain level and to take into account any mastering overlays, pseudo-VCA subgroups, etc.: what coefficient values to create for a parametric EQ section changed to differing parameters of frequency, level, and *Q*. In addition to being a fraught exercise in database management, there's some pretty good math in there too. (In DSP software design, one strives to keep the actual DSP algorithms as straighforward [fast] as possible, leaving as much squirrely and calculation-extensive stuff as possible to the coe-gen code in the host.)

Since a major part of the thrust toward digital consoles has been their promise of storage and recall, statically (snapshot) or dynamically (as in real-time automation), it is beholden to the host to manage the data transfers involved. Everything that may need to be stored is already in the host, but the software routines and hardware to facilitate storage/recall need to be present. A console can be quite self-contained in this regard if the data set is relatively small; on-board flash memory may suffice. Otherwise, whirling and whining hard drives may well be necessary. In the event the console is integrated reasonably closely with an audio recorder, hard disk, or otherwise, the automation data may get squirted onto that as a sideband.

In a split console (control surface is separate from the guts), the host also has to manage intercommunication with the control surface; typically this is done with an Ethernet variant, which demands the existence of a TCP/IP stack for the communication protocol and hardware to terminate the Ethernet.

25.24 Digital Audio Workstations (DAWs)

Fig. 25-151 shows the major (usually indivisible) elements within a digital console system and their relation to each other.

User Surface. This has indication of control positions, metering, and means of controlling the audio processing and can range all the way from the sea-of-knobs large-format console-style surface to graphics on a PC screen with a mouse.

Surface Host. A micro to look at, make sense of the controls, and drive the metering. This can vary from being a small embedded micro to a large PC-like processor, depending on the size of the surface and if it is expected to do high-speed communication, should the control surface be remote. It can also not exist, its nonetheless necessary functions being subsumed by a PC's CPU processor.

Processing Host. Takes care of looking at the data being passed to it from the surface host, creating the necessary coefficients for the audio processor, and also looks at the various and many metering returns from the audio processor, rendering them down into a form

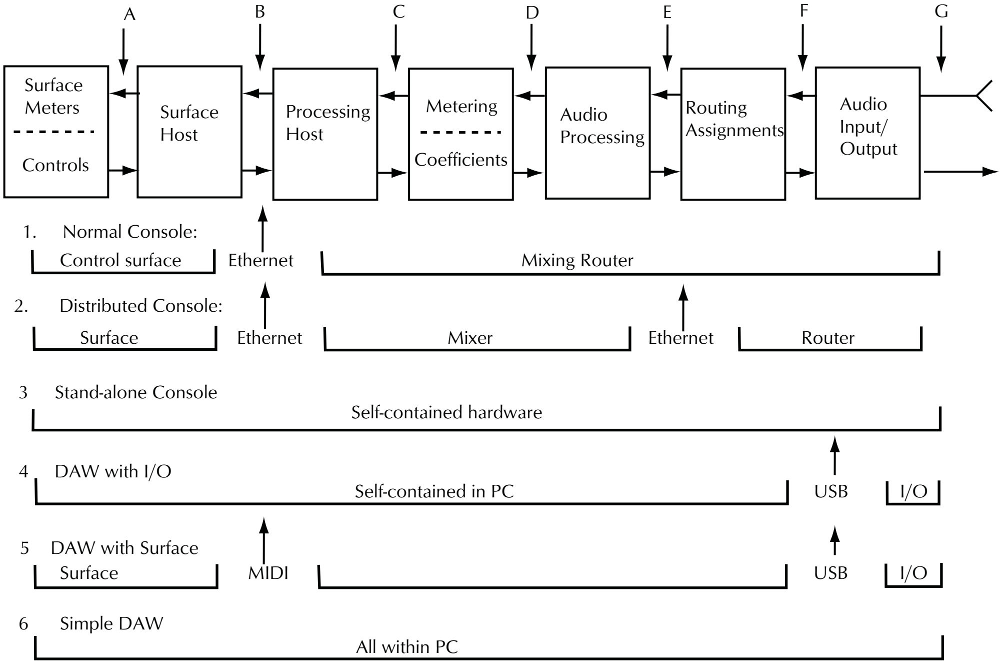

A
B
C
D
E
F
G
Surface Meters
Controls
Surface Host
Processing Host
Metering
Coefficients
Audio Processing
Routing Assignments
Audio Input/ Output
1. Normal Console:
Control surface
Ethernet
Mixing Router
2. Distributed Console:
Surface
Ethernet
Mixer
Ethernet
Router
3 Stand-alone Console
Self-contained hardware
4 DAW with I/O
Self-contained in PC
USB
I/O
5 DAW with Surface
Surface
MIDI
USB
I/O
6 Simple DAW
All within PC

Figure 25-151. Process distribution in various styles of digital mixers.

useful for the control surface to display. This is usually a sizable and capable processor, unless its functions are being done by a PC's CPU.

Audio Processing. This uses the coefficients from the processing host to modify the audio path(s) as desired, usually within a raft of DSPs, either aided by or supplanted by FPGA's, unless, of course, the audio processing is within a PC's CPU.

Routing Assignment. Decides what audio path is going to get routed through what path within the audio processing and so what controls apply to it. Also decides on what output ports processed audio appears. Often an additional soft matrix done in a DSP mixer (particularly when integrated with Audio Processing), or a hard switch in FPGA, and often a stand-alone product but entirely within the capabilities of a PC's CPU.

Audio Input Output. This is the termination of audio sources and destinations and their conversion into a form of digital information the Audio Processing can assimilate. Examples are A/D and D/A converters, usually many, and SRCs (Sample Rate converters) to seamlessly integrate external digital audio sources. Often stand-alone, in cages dedicated to their purpose, locally or separated from the rest of the system. Often integrated with the Routing/Assignment, and sometimes along with audio processing. And sometimes just plugged straight into a PC. (Bet you can't guess where this is going…)

The paths between any of these blocks may be broken and subject to transport if need be, but it is far more likely by way of practicality at labeled junctures B, E, and F; for instance, it is generally easier to transport the rendered, lower data concentrations of control parameters and meter data between the Surface and processing hosts at (B) than it would be to try to move the raw preprocessed metering data and coefficient sets, as would be the case if a split were made at juncture (D).

Nearly any audio console with digital control fits into this loose model and contains all these elements in some form or other; even a DCA (digital control of analog) console, with a control surface separated from the processing electronics at interface B, is similar to case (1), a normal console. So it can be seen that the lineage from analog consoles, through DCA, pure digital consoles (which at least initially mirrored analog consoles almost totally), through to DAWs is quite plain.

The subdivisions of these processing blocks for different classes of console, and with likely transport means, are shown in Fig. 25-151:

1. A normal control-surface/mixing router arrangement is shown with a single major transport requirement at juncture (B).
2. A distributed mixing-on-the-network style console is shown with network links inserted at (B) and (E).
3. An all-in-one box arrangement (emulatingo-stand-alone analog consoles and simpler digital consoles) is shown with no transport insertions.
4. A DAW with an external input output unit has a connection at E.
5. A similar DAW but with an add-on physical control surface has links at B as well as E.
6. A simple DAW with limited I/O output having no need of external interconnection.

This latter uses the host PC's GUI for control and display, the PC's CPU to do all the hosting and audio processing, and internal converters to get the audio in and out. A surprise may be the extensive use of semi-pro or domestic communications schemes in the DAW contexts —for instance, MIDI (musical instrument digital interface) for the control surface interconnection, and USB as the audio transport to the audio I/O interface box.

The major underlying message from all this is that DAWs are consoles, too! In broad-brush architecture as shown in Fig. 25-151 they are—since they have to perform all the same functions—indistinguishable from "real" consoles, which is actually an understatement, since in many respects DAWs are more versatile and powerful.

25.24.1 The PC

A decently fast and capable central processor(s); a reasonably easily crafted and programmed graphical and user interface; fast, inexpensive, and capacious memory; and omnipresence all afford the PC an enviable basis for audio production. It is and nearly always has been a more cost-effective platform than any purpose-built digital audio system of comparable facility. All the technical advantages made it a natural basis for initially fairly elementary audio functions such as a hard disk recorder/stereo editor, up to today where entire multitrack recording/editing/processing systems readily fit on a laptop—the like, of which would have been the envy of major studios just a couple of decades ago. Despite the best efforts of operating system manufacturers to make real-time audio streaming into and out of PCs problematic, the PC is a formidable tool.

25.24.2 MIDI Sequencing—Where It Began

An early PC application was in the recording, storage, manipulation, and automation of MIDI-encoded musical parts, to facilitate the assemblage of songs. This did not involve any audio, per se, merely the management of streams of MIDI commands against time. These were then issued in sequence down a MIDI path to attached music synthesizers that played the music itself.

The desired ability to compose, rearrange, copy, and time-slip parts in relation to others in synchronization gave birth to extensive and powerful automation, which largely outshone concurrent traditional console automation schemes.

Recording and manipulating audio on a PC occurred when processor speed and disk drive size and access speed allowed (two-track editing became commonplace, resulting in stereo tape recorders plummeting from hallowed possessions to doorstops virtually overnight). Although the means of getting multiple simultaneous live audio streams into the systems lagged, it was certainly possible for multiple tracks to be recorded sequentially so building up a true multitrack recording, and this was exactly the mode of operation prevalent in basement studios anyway.

And so it was not the least bit surprising that the major exponents of MIDI sequencing software became the major exponents of PC-as-studio, and their approaches from MIDI world translated over into audio world reasonably well, despite significant differences in philosophy. This does explain why those previously steeped in traditional recording find the assumptions, methods of control, and even terminology of sequencer-studio tools quite alien, while those who have grown up with it regard traditional techniques (and terminology and assumptions) to be, well, odd and quaint. MIDI sequencers have cast a long shadow over today's audio processing.

Many of the strengths of the sequencer applied to audio readily in ways unthinkable before—time-slipping or copying individual tracks or segments, unlimited takes of tracks or segments being treated as related parts rather than completely separate tracks, as examples. As the recording hardware (PC) became more powerful and the number of instantaneously available tracks increased, a deliciously ironic approach has come to the fore: originally, the sequencer shuffled MIDI elements around, in the hybrid audio-plus-MIDI the two were treated in parallel yet separately, but now it is common for all the MIDI tracks to be rendered as audio onto audio tracks just like live sources, and the audio control and automation methods rule.

25.24.3 DAW Audio Ins and Outs

Means of moving audio around are covered in more detail in Section 25.25. DAWs just like any other console need to get audio in and out, and from the lesser to the greater this can include:

- PC's built-in sound card. (It had to be mentioned, and besides, who honestly has never used one in a pinch?) Typically analog-in, analog-out (sometimes S/Pdif), at very low domestic signal levels, and of generally indifferent to awful quality. But convenient.
- USB/Firewire. Links to external sound card convertor boxes from stereo in/out up to as many as 16 in/out, see Fig.25-152.
- ADAT. 8-in or 8-out via fiber-optic cable.
- MADI. Up to 64 ins or outs via coax.
- Ethernet. Either true TCP/IP Ethernet or audio-specific UDP variants using the same hardware typically 64 I/O for UDP.

Specific drivers—and in the case of ethernet and variants whole suites of interface code—need to be installed on the PC to deal with the audio on these various schemes. The DAW software has both input and output routers that can pick which incoming sample within a stream goes to what input, thus track, and which DAW output gets sent out what slot.

So far, the PC-based, sequencer-modeled audio control approach looks like a multitrack recorder (of virtually unlimited scope) with wicked automation and editing. But what of console-style audio processing?

25.24.4 DAW Internal Audio Paths

The audio routing within a DAW tends by design to be quite basic: Fig. 25-152 shows this as being essentially a route from input to a recorder track, thence from either before or after the track to a mix bus (or buses), thence to an output(s). Recognizable console like features such as a fader and panning are included just to show a typical starting environment.

What all the Xs mean is that it is possible to "drop in a plug-in" (translated: apply an instance of a signal-processing software module) or apply the signal at that point to anywhere else there's an X. An output bus, for example, can and frequently does get routed back to be a recording track source (bouncing in oldspeak); many modules may be inserted concatenatedly at each X. There is considerable flexibility. This approach—nearly everywhere being an insert point and only providing access for processing—as opposed to the "everything's in there in case" traditional console model allows what processing power there may be to be applied as and where it is needed while leaving all other paths unfettered.

25.24.5 Plug-Ins

A plug In is a collection, library, of disparate software programs that variously (a) actually process (or generate) the audio in some form or other—e.g. EQ, dynamics, delay, reverb, etc., or a MIDI musical instrument; (b) provide a graphical module for display on the system's GUI, replete with knobs, buttons, dials, gauges, meters, and blinky-lights, (c) calculate the conversion of the parameters from those controls into coefficients that the actual signal processing can understand; and a render metering data in the reverse direction, from audio to GUI. All the "handles" typically become available to the system's automation system, either directly or by being MIDI addressable—in other words, the module can look to the system as just yet another MIDI slave device, which can automate it accordingly.

Standards have evolved (in the form of requirements by major DAW players) for plug-ins; VST (virtual studio

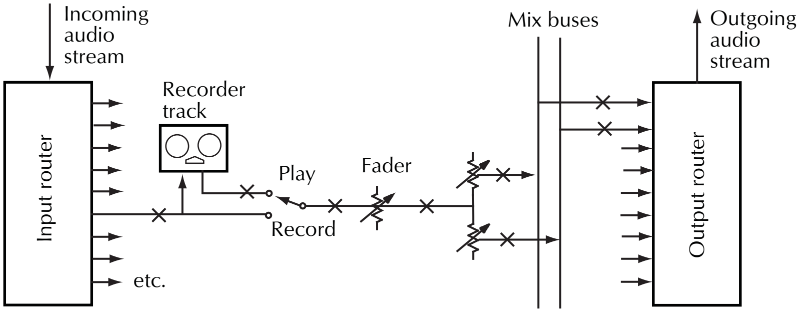


Figure 25-152. Simplistic DAW audio path.

technology) and DirectX stand out as the more popular nonproprietary schemes affording wide interchangeability between different flavors of DAW.

As mentioned, many plug ins are MIDI musical instruments or devices in their own right, but the widest variety is in audio processing. Most DAWs come with a decent suite of generic modules that allow all the traditional functions, plus some others that had hitherto been rack-box fare, such as reverberation units, flangers, etc.

There is a huge variety of modules available, some being specialized, and many that emulate, with greater or lesser degrees of success, existing real-world boxes, either contemporary or classic. These vary wildly, from being merely a pretty face (GUI) controlling a set of disappointingly cookbook algorithms and being passed off as something special, to exceptionally and painstakingly crafted emulations of existing products, accurate even down to little-known quirks. Emulations aside, DAWs have reached such a level of acceptance and usage that there are module manufacturers for whom plug-ins are their sole business.

25.24.6 DAW Limitations

Any description of DAW limitations is doomed to become laughable as their underlying power inexorably increases. A solution to the lack of signal-processing horsepower in the earlier days was to offload the audio signal processing onto DSP farms, either on slot-in cards that fit within the PC's box itself or in an external frame. This afforded far superior overall performance than was then possible from the PCs CPU alone and is still an approach taken by DAWs aimed at the professional area. The downside is that it can lend itself to creating a proprietary technological island, exacerbated by the use of nonstandard fileformats, making interchange between other types of systems difficult.

Clever approaches to make the best use of limited processor steam revolve around using otherwise dead time and the almost limitless ability to store recorded tracks. As an example, if an EQ is applied to a track, rather than run that EQ in real time each time, it's played (along with possibly dozens of others, which may very well drown the system), it is run very quickly and quietly, once, across the length of the track, which is then saved as another track. That way the system just plays back a pre-EQed track rather than having to run an EQ—a huge saving in resources. If a change is made to the EQ halfway through a playback, the EQ runs in realtime from the change but at the end of the playback the resultant overall EQed track is saved as yet another track. The system keeps track of which track is the most current: this is also key to how DAWs can seem to have boundless ability to roll back or Undo changes—in addition to the automation remembering all the changes, all the older tracks are still available for instant application. Effects tracks, reverb passes, etc. need only be striped once, and never need to eat PC power again.

Reference to Fig. 25-152 shows that it is, theoretically, possible to avoid the use of the recorder altogether and simply use the DAW as a straightforward mixer. However, one has to remember two things:

1. Every instance of every plug-in will use juice, and one will sooner rather than later find out how many is too many—the resource mitigation dodge of prestriping effects on tracks doesn't work mixing in real-time, where everything has to be happening at once. Large-scale live recordings can be done on such a DAW—the sources will all go straight to track with little on-the-fly processing being necessary, and it can all get fixed in the mix.
2. There may be excessive latency (input-output delay), mostly from the acts of getting the audio into and out of the box; this may well be audible or annoying in some circumstances.

25.25 Moving Digital Audio Around

As is plain from the earlier discussion of digital audio mixing and processing systems, and in particular that there are few constraints on where the constituent bits are physically in relation to each other, there can be an awful lot of audio to shuffle around between them. The term intra-console is used to distinguish this interconnection between bits of a console system, as opposed to moving lumps of audio around in a facility. Often, though, this gets blurry!

Sometimes all that is required is the movement of some audio from one place to the other, but increasingly there is a requirement to have all—or some—sources available at all—or some—destinations in a free-grouping arrangement. This takes on a life of its own as a network. Most end-to-end signaling types as described here can be made to become part of a such a network if they are arranged to have one of their ends terminated in a hub, in star configuration with other end-to-end links; the hub—router—has the intelligence to route the signals telephone-exchange-like accordingly. Some described transport mechanisms are designed to be networks, or network like, in their own right—the 800 lb. gorilla in this world being Ethernet.

25.25.1 Moving Audio—small-scale

25.25.1.1 AES-3

Stereo pairs (or pairs of monos) have been catered for by the venerable AES-3 standard; this is a Manchester encoded stream of two (up to) 32 bit audio words, some informational tags for link status and format, and a number of user bits that may be used for anything from turning stuff on and off remotely to serially carrying metadata (program-specific information) or more complex real-time control data. It was designed to be robust, simple, and as usable as possible in the predominantly analog world into which it was born, even down to using the familiar 3 pin XLR connectors in its usual implementation. With minor updates, mostly concerning connection variants and data rate (it now handles the once-unthinkable 96 kHz with ease), it still serves well. It is a very close cousin (indeed the underpinnings are chemically indistinguishable, and use the same chip sets) to the domestic S/Pdif (Sony/Philips Digital InterFace). The audio is treated identically, but the format and informational tags differ. It is common for AES-3/-S/PDIF receivers to be set up to strip these off so as to allow universal connection, but obviously this is at the expense of any metadata that may accompany the audio stream, and, if this is of the least concern, any digital rights mismanagement flags.

A performance downside to AES-3, particularly with early implementations, was recovered clock jitter. Best performance is achieved by reclocking at the receive end, either by SRC (Sample Rate Conversion) or the use of very good flywheel phase-lock loops to reestablish solid, quiet clocking. If the facility is homogenous with everything running off a master clock this is less important; "bits is bits" and as long as they arrive within the same framing period (e.g., 20.8 µs at 48 kHz) and sample clock period, and any D/A is done with the same pristine clocks as any A/D, transmission jitter is irrelevant.

25.25.1.2 AES-42

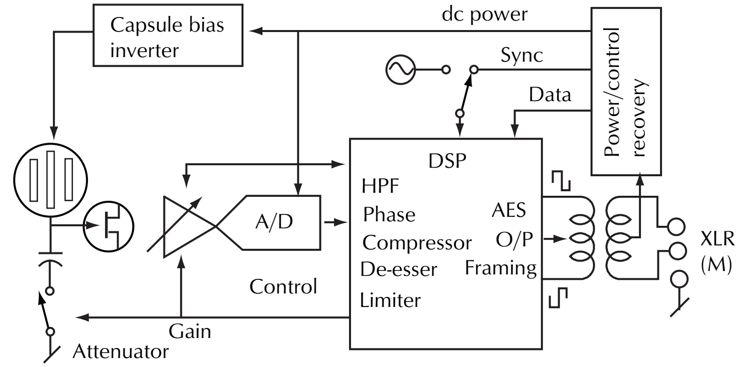
As will been seen in the later mention of USB microphones, there is a drive to push digital as close to the source as possible, in that case for simplicity's sake, in proaudio for performance. The concept of putting mic preamplifier, A/D converter, and processing inside the microphone itself is at one and the same time seductive and puzzling. The idea of simply taking a digital stream (possibly in AES-3 format) straight from a microphone into a digital system holds strong sway; reflection shows that this—in any meaningful system—means either the addition of a plethora of hitherto unknown knobs and switches on the microphone itself or the means of remote-controlling all those functions and takes the shine off the idea somewhat, particularly to those to whom a microphone is something one simply plugs in and uses.

As has been made clear, there is little that binds a particular function to a particular physical location or piece of system hardware or software. Given that, some mouse-and-screen GUI widgets to control the microphone parameters, or indeed a physical set of hardware knobs and switches to do the same, don't care whether the target is in the same box, another processor, or even on the same continent. That, in this instance, the target is on the top of a shiny microphone stand in the studio is irrelevant. So, not only is a means of getting digital audio from the microphone necessary, but means of getting the control parameters or coefficients up to the microphone, as well as a synchronizing reference clock so that the microphone's pristine audio doesn't have to suffer the immediate indignity of a sample-rate conversion to match the rest of the system. And, of course, a means of powering all this.

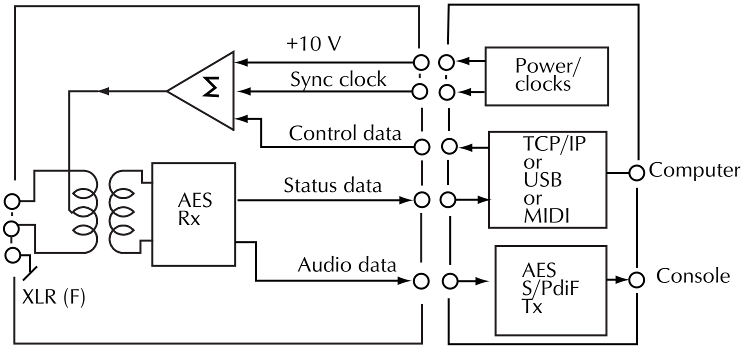
And so was born AES-42, in an effort to standardize all this before multiple incompatible approaches dissipated the concept's appeal. Fig. 25-153 shows in outline form its scope.

Many hitherto console functions have found their way into microphone control via AES-42. Although the scheme is not limited to these, the Neumann TLM-103-D digital microphone, for example, allows gain, microphone pattern, absolute phase, high-pass filter, an in-built compressor/limiter/de-esser, and a peak limiter's parameters to be controlled. It's easy to see where that's headed; no need for console channels as we've known them.

The normal connectorization is via the old familiar XLR, although the XLD is suggested for circumstances where confusion with other XLR-using systems could potentially result in damage. As would be expected, signal formatting owing much to the familiar AES-3 is used to retrieve the audio, which ordinarily comes differentially down a shielded pair; user bits in the data stream relay fixed data such as the microphone's manufacturer, model number, and available controls; variable data such as instantaneous parameter value are also available by this means. Now the fun begins—power is sent phantom style (common mode and with reference to the shield) back up the line; instead of its merely being regulated down to power the microphone and its electronics, it is also modulated with control data and a synchronizing word clock, which are filtered off and


Capsule bias inverter
dc power
Sync
Data
Power/control recovery
DSP
HPF
Phase
Compressor
De-esser
Limiter
AES O/P Framing
A/D
Control
Gain
Attenuator
XLR (M)

A. An example AES-42 digital microphone.


+10 V
Sync clock
Σ
Control data
Status data
Audio data
AES Rx
XLR (F)
Power/ clocks
TCP/IP or USB or MIDI
Computer
AES S/PdiF Tx
Console

B. Basic AES-42 host interface C. Option example.

Figure 25-153. AES-42.

used to instruct the microphone's processing. The microphone's sample rate may either free run, in which case it is a master (but will probably need SRCing to work in a system of any complexity), or it can be slaved to the synchronizing word clock. The latter is favored, if available.

Present digital microphones using AES-42 have a choice of termination, depending on whether the system into which it is plugged already speaks AES-42 in that particular microphone's dialect (and so control from that system is implicit), or via an external interface box which permits a computer running the appropriate and proprietary control software to talk to the microphone and audio data recognizable as AES-3 or S/Pdif stripped off for use.

An interesting side note is that Neumann, a major influence over the scheme and early adopter, make claims that such an arrangement results in better overall dynamic range than traditional microphone connections. The premise is that conversion of the capsule audio down to the common low-level microphone interconnection standard of 150 Ω/sent through a wire/low-noise amplification/then conversion of that signal to an A/D convertor within the console, is intrinsically noisier than the more direct connection of the capsule with optimum impedance transfer to the convertor within the microphone itself, without intervening transformations and stages. Their claims of convertor performance so used are impressive.

25.25.2 Moving Audio—Multiple Paths

More than a stereo pair calls for more radical answers, and as is typical with fast-moving development tended to outstrip standards-making—never mind the commercial impetus to try to capture users within a proprietary format. Two formats, one from pro audio, the other from semi pro, stand out from the earlier days of multi-track recorder/console interconnection:

25.25.2.1 AES-10—MADI

This format is very common for the interconnection of digital reel-to-reel recorders (whoever thought we'd be weeping nostalgic for those?) and older large-format digital consoles. It carried up to 56 audio channels (now 64) originally down inexpensive TV-style 75 Ω coax (plenty enough for a 48 track recorder) and owed a lot to FDDI, an older communications network backbone format. Being unidirectional meant that a MADI link in each direction to a recorder was necessary. Latency was quite low, and the availability of chipsets made implementation fairly straightforward. An oldie-but-goodie, it is still used to an extent in the pro audio world by some manufacturers for overall system interconnection and intra console knitting.

25.25.2.2 ADAT

ADAT is a simple (both in hardware and signal format) unidirectional fiber-optic interconnection, originally, to get 8 audio signals into or out of the once highly popular Alesis ADAT VCR-based 8-track recorders (which can be thanked as being the likely tipping point of recording from uptown to basement). It is still an interconnect of choice in semi-pro recording equipment, where "pieces of eight" is adequate or sensible, as when additional functionality is marketed in such a modular fashion.

Being a strictly hardware interface it is wholly deterministic (audio arrives exactly when expected) and with very low latency. Although inexpensive chipsets are available, the format lends itself to low-impact implementation in (possibly already existing) FPGAs (Field-Programmable Gate Arrays) within a product design, so incurring near-zero add-in cost.

An ADAT frame, which can carry up to eight 24 bit audio words, is 256 bits long at a clock rate of 12.288 MHz for 48 kHz. There is a 16 bit preamble containing a 10 bit frame-sync period and four user bits for control/messaging. (The arithmetically astute will wonder where the other 46 bits went; they are used throughout the frame after every 4 bit nibble—except in the frame-sync period—as synchronization zero-value bits). The bits are scrambled (Manchester-encoded) to non-return-to-zero to remove any tendency to have a dc component. Some of this—in particular, the syncing and NRZ—had a lot to do with coping with the vagaries of VCR tape transports, but as a long-standing standard with millions of installed instances it holds up very well and doesn't warrant the potential confusion revisitation and redesign would incur. It is hard to envisage a simpler robust multichannel self-clocking interface, and its designers deserve full credit!

25.25.2.3 USB

A somewhat surprising development has been the adoption of the humble USB connectivity of PCs to move moderate amounts of audio about. This reflects the massive shift over recent years from the large-scale studio-as-shrine approach of the recording business to small-scale home or demo studio recording becoming the new mainstream.

Conceived as a replacement and expansion for RS-232 serial connections (and related mouse/keyboard interfaces) for PCs, the early USB implementation (e.g., v1.1) was hard pressed to reliably move a stereo pair of 44.1 kHz about, but the upgraded USB2 with its nominal 480 MHz data rate changed all that. As an example, Fig. 25-154 shows a 1U rack-unit box by Tascam (beneath the laptop, above the mega mic preamps) that readily simultaneously transports 16 audio paths to, and 4 back from, a PC running DAW software, all via USB2. USB2 seems to have eclipsed Fire Wire (IEEE 1394), a similar-speed (if network-capable) interconnection that hitherto briefly reigned in the sphere of small-scale PC audio transport to external A/D and D/A boxes and such.

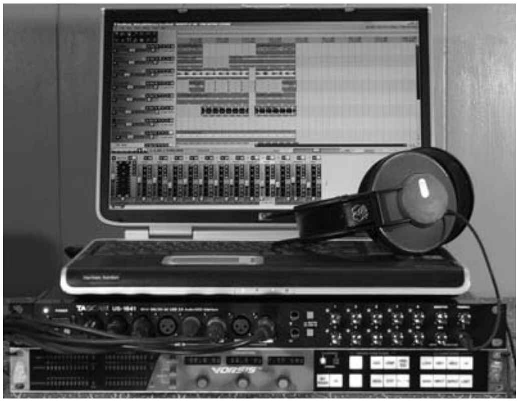


Figure 25-154. A modest-sized DAW running on a USB link between the audio interface unit (center) and the laptop. The (almost free) DAW software allows for 48 simultaneous recorded tracks, with significant audio processing. Complete with shown external mic-amps and computer, the cost of this outfit is about that of a decent microphone; its 1990 equivalent in facilities and performance would have cost the same as a decent car, while the 1970 version would have equated a decent house, which would have been needed to fit it all in, too.

It is not at all uncommon for small mixers— analog or digital—to present their outputs and accept a returning pair of inputs via USB; small hand-held recorders likewise; microphone preamplifiers; even

USB microphones, which only work in that context. The PC to which they are connected can recognize such simple interconnections as just an external Sound Card, and has generic drivers built in to cope. Zero-effort connectivity. Better performance and more advanced control and features can be achieved with special drivers installed in the PC, but the instant connectivity thing is hard to beat.

One rightly has to be circumspect about jitter performance on a transport mechanism that certainly was never characterized for digital audio streaming with very stringent clock recovery requirements; in the case of the small DAW setup described, since A/D and D/A are in the same box, and assuming the clocking is done conscientiously within, the overall performance is limited by that, any vagaries of the link, its latency, or computer timing being irrelevant as long as the clocks remain synchronous with the data. Such are the dangers of any transport scheme where the clock is solely implicit to the data, with no external reference.

Of far more concern, however, is the computer's operating system's handling of audio, which can make a land of horrors transcending any worries about link jitter; this is typically addressed by the loading of unit-specific drivers (ASIO in this case) into the computer, which blow right by the operating system's clunky hardware abstraction scheme, and instituting delay buffers capable of absorbing most temporal irregularities. Nevertheless, a USB (and perhaps more so FireWire) link's performance is often dropout-limited by the host PC's handling of (deferred procedure calls), the stacking up of interrupts and time-related routine calls that take longer to address and clear than the link's buffers can sustain. Sometimes a lot of effort has to be put into disabling features/programs/peripherals (like wireless networking in particular), updating or finding the right/better drivers, optimizing this and that, installing replacement hardware, and general hair tearing, just to get a PC to adequately pass/process meaningful amounts of audio. The PC really is not a shining beacon of a streaming-audio-friendly environment!

25.25.3 Digital Audio Networking

25.25.3.1 CAT-5/RJ-45 interconnectivity types

The following communication schemes all typically use the widely available (even from the local stationery store) networking style cabling, exemplified by CAT-5 or CAT-6 cable terminated in the little plastic RJ-45 phone like connectors, and indeed often share the same terminating MAC and PHY electronics. What actually goes through them can differ wildly though. As will become plain, this enabling technology has also expanded the notions of what can be done in the context of moving large amounts of audio around, blurring the distinctions of transport, mixing, and processing.

25.25.3.2 TCP/IP—Audio over Ethernet

Ethernet, using its handmaiden TCP/IP platform, is highly popular—ubiquitous—and there is a large base of skill in operating and maintaining networks based on it. This has lent impetus to trying to use it for things for which it wasn't really intended and is not particularly apt, such as moving professional audio. Immense effort and marketing has gone into making it work adequately. Probably best placed elsewhere.

Any Internet user knows how facile it is to move audio around, either in chunks as files or drip fed as streaming, either on a local network or the internet itself. A seemingly sensible follow-on would be to wonder if the self-evidently already existent methodology could be used for moving large quantities of audio around in a digital audio network.

Well, the short answer is "Yes." Given the existence of Ethernet connectivity and a good IP stack (network operating firmware) in each of the required connected units, a significant number of uncompressed audio channels can be moved around successfully by such a network. However, the long answer begins. Aspects of its performance at best are eclipsed by alternative methods. The only real advantage is the previously mentioned wealth of user familiarity with, and labor skilled in, TCP/IP networking. There are significant drawbacks.

In the Internet example, the audio is almost always compressed by MP3, WMA, AAC, Ogg-Vorbis, or whatever format to radically reduce file size or to fit within a required streaming rate. Except for a few arguable examples such as news, remote, or commercial distribution for radio broadcast, compressed audio has no place in professional audio. So suddenly the necessity of uncompressed audio payload can be ten times or beyond the size of domestic audio. Network congestion effects loom that much closer, that much sooner.

TCP/IP is a packetized system and incurs at best a minimum packet assembly/disassembly time at the ends in addition to the relatively quick transmission times. The packets are ordinarily comparatively small—in a streaming sense—multiplying the processing/deprocessing overhead and bandwidth wasted in sending packet headers. This can be tweaked, however.

This all incurs latency (i.e., a delay between input and output) that may or may not be acceptable. Live applications—say, broadcast or sound reinforcement—might well have issues, particularly if many links' latencies become cumulative. These are relatively minor latencies, though, compared to what's to come in a real environment. Oh, yes. It gets worse….

Congestion is paradoxically key to TCP/IP's main limitation for audio, considering it was designed to—and does—handle congestion superbly for its intended use. In the absence of any other network traffic that may contend with a primary audio stream, the packetized audio will likely arrive unmolested and in order, and a fairly high density (lots of audio) may be passed from a point A to a point B. In short, in a point-to-point dedicated link, audio via TCP/IP can work reasonably well.

Unfortunately, that's not what the network concept promises: multiple independent streams from multiple sources and with multiple destinations sharing the same wire infrastructure. As soon as other traffic hits the network—say, another audio stream from point C to point D—try as the carrier-detect collision avoidance mechanisms inherent to Ethernet might, packets from one stream will unavoidably tread on those from the other. One of TCP/IP's great strengths is that it recognizes such events and deals with them handsomely; each stream gets the opportunity to resend its broken or unacknowledged (i.e., lost) packets, and the receive stack knows to reassemble the stream in the correct order from the now possibly out-of-order and certainly delayed packets. So what's wrong with that?

25.25.3.2.1 Buffering Latency

The network has lost any tenuous claim to determinism —predictability—it may have shown, since collisions and recovery therefrom are unpredictable both in frequency and recovery time. Determinism is, in short, knowing exactly and consistently when recovered audio is ready for use—absolutely essential for streaming-type or real-time audio or dropouts occur. A pure, isolated, low-density point-to-point TCP/IP link can be close to deterministic and with a relatively short latency, predictable from the above-mentioned packetization, framing, and transmission times. Even so, it's only close—other traffic still exists on the link, in the form of ACK (acknowledge) replies for each sent packet: Yes, collisions can occur between the real data and its own ACKs! Real-world, where multiple paths on the network are in use, significant collision-recovery times get thrown into the mix and this now unknown added time becoming even more so and approximately geometrically longer as the amount of traffic increases. There is also the very strong likelihood, nay—certainty, that packets that have been stepped on and repeated will arrive out of sequence, the repeats only getting through sometimes many frames after several in-sequence packets have progressed.

The workaround—trading a fixed, known, longer latency for a shorter but unusably unpredictable one —is by instituting a fifo buffer (first in, first out) at each receive point. In order to allow time for packets to be eventually received and juggled back into order, this fixed deliberate buffer latency has to be incurred; the more traffic, the more latency is required, and on a busy network this can be in the tens or hundreds of milliseconds to encompass worst-case congestion effects. Although acceptable in some circumstances this is difficult to swallow for many audio applications—particularly those where there is a requirement for humans to listen to themselves live through such a system. Consequently, lowish-latency and pseudo-deterministic audio links are usually recommended to be placed on discrete one-to-one links with little risk of contention. Which really rather begs the rationale behind using TCP/IP, and the promise of "networking" upon it. Oh, well. All of these ills are exacerbated if any other traffic is permitted on the same network—which finishes off the naive notion of running significant amounts of audio on an existing office network. Worse yet is expecting sensible behavior if incoming or outgoing real-time audio is expected through the Internet—build-out latencies may have to be far, far longer to absorb the hairiness of the unknown out there! Again, this may be acceptable in some circumstances—after all, if one is using the Internet, the likelihood is that the audio is going a long way away, where no frame of reference in time exists to its source.

One saving grace of the general move to gigahertz Ethernet (as opposed to the more commonplace 100 MHz variety) is that everything happens much quicker, and that for normal practicable amounts of traffic the collision rate and recovery times go right down and so the build-out buffering latency can be radically reduced; TCP/IP as the basis for an audio network reaches a lot closer to the promise, as opposed to the highly marginal on-the-edge behavior of any meaningful size system at 100 MHz. The advantage is not so much that ten times the traffic could theoretically be handled, but that a similar amount of traffic can be handled well; latencies in the single-digit milliseconds are readily achievable, which, if not too many passes through the system are attempted (remember, the per-pass latencies add up), is a generally acceptable performance.

25.25.3.2.2 Latency—How Much Is Too Much?

Despite many learned researchers' effort, most data concerning the audibility of latency is based on the anecdotal and apocryphal. But there is no substitute for being on the wrong end of a broadcast presenter ripping off his headphones and spewing invective as establishment of an incontrovertible benchmark.

We won't even discuss delays that are long enough to be discernible as a delay, or a discrete echo; that is obviously way too long, and everyone, trained or not, has a hard time speaking normally when fed such into headphones or monitors. No, it's that mushy area less than, say, 50 ms delay—a period of time below which the ear/brain attempts to integrate all correlated sources into one—that is of concern.

Latency is an issue where a performer is listening directly to a delayed version of him or herself; two situations to keep in mind are a DJ wearing headphones or a stage performer with in-ear or conventional floor/side-fill monitors. An important thing to note is that very different answers from these people as to what is noticeable, annoying, or untenable are garnered depending on whether they are introduced cold to a system with delay, or are steadily introduced to it, particularly in the cases of headphones/in-ears.

Talking, one hears oneself not only by what's coming through the headphones, if they're open-frame headphones (i.e., not enclosed), by room spill, but also by bone conduction within one's own head. This latter is distinctly band limited, and what is passed is usually just the fundamental and possibly early harmonics of vowel sounds. Interference between this and what is being stuck in the ear causes a nonflat perceived frequency response, with cancellation notches and corresponding reinforcement summations. (It is the same mechanism as the audio effect *flanging*.) This is in general no real problem—one quickly accepts that sound as being normal, the sound of oneself wearing headphones. Deliberately introducing a different delay by even only a millisecond or two is immediately perceptible—the interference cancellations/summations change—the sound changes. This is why many tests attempting to establish acceptable latency by steadily increasing delay have resulted in unrealistically low values; the relative changes in coloration with even small changes in delay are very easy to perceive, even by the unskilled—and immediately flagged as a problem.

Conversely, if one were to present a subject with a delayed headphone feed even quite a bit larger than this (without previously having had chance to establish a reference), the interference-related sound would readily be accepted as normal.

In daily use on countless radio stations are air chain processors with delays in the 10–15 ms region; this, in addition to other latencies in the loop path from microphone to headphones listening off-air, means delays approaching 20 ms are commonplace and to a greater or lesser degree, accepted. Much more than that, though, engenders complaints of the sound being disconnected or hollow and distracting.

Time-alignment experiments conducted on large-scale rock'n'roll sound systems reached broadly similar results; 20 ms monitor delay was as much as could be tolerated by most performers, although some could detect far less, but most readily acceded not to be too bothered by it. Delay between the performer and the PA, particularly in a large venue, proves relatively unimportant for two reasons: firstly, the performer has much more present (louder) monitoring to which he's likely paying much more attention, and, secondly what scatters back from the PA is quite diffuse and decorrelated anyway. In all cases, the threshold of unacceptability is very crisp—definitely a straw-that-breaks-the-camel's-back situation.

The main thing to be considered in all this is that latencies add: each pass of a signal through a signal link or network; each piece of gear or processing to which it is subjected; each propagation delay adds up to often be significantly bigger than one might expect. Just one more teentsy-weensy little few link milliseconds through a TCP/IP pipe might just break it.

25.25.3.3 UDP

UDP—*User Defined Protocol*—essentially uses the same (fabulously inexpensive and readily available) Ethernet-style connectivity, hardware, and chip sets but with a far simpler messaging protocol than TCP/IP and better suited to the application at hand. It is then of no surprise that the majority of wide (more than two paths) commercially available audio transports use a UDP variant. One hundred MHz Ethernet hardware using UDP can afford very low latency and wholly deterministic audio paths, with, for example, typically 64 discrete paths bidirectionally at 48 kHz sample rate. One GHz hardware/firmware allows correspondingly greater capacity.

As mentioned, most manufacturers' audio transports use this mechanism or something like it; there are countless varieties, all proprietary and utterly incompatible, of course, since there is a strong commercial impulse to keep everything in house. A standard, AES-50, attempts to bring sanity and compatibility into the gigahertz realm, with the myriad 100 MHz schemes already considered a lost cause.

At its simplest, a big packet consisting of a header and then however many audio words of whatever width is constructed, then sent down a dedicated Ethernet hardware circuit, every sample period; at the receive end the simple format is readily decoded and the constituent samples recovered. It is a reasonable assumption that on a dedicated line the packet will be unimpeded and arrive intact, and such links typically run raw with no mechanism for error trapping. Of course, it is entirely possible to build in error detection and correction in case bits get hurt somewhere. It would have to be a fairly weak link for this to get exercised much, and such mechanisms raise the bugaboo of TCP/IP systems—building out a fixed latency to allow for the randomness of the errors. In short, these systems run just fine without and typically do.

Audio is only part of the whole picture. Metadata accompanying it, logic switching contingent on control, control data and metering data all have to be considered and accommodated within the link for it to be a fully usable system in any meaningful context.

Such UDP links are typically bidirectional (but sometimes unidirectional) end-to-end closed links. In and of themselves, they don't constitute a network, which can loosely be described as *anything to anywhere*—any source connected to the network may be picked up by any destination. There are two general schemes for turning these one-to-one links into networks: cascading them node-to-node, with a modified signal passing along each link (Serial), or arranging them all to radiate from a central hub (Star).

25.25.3.3.1 Serial or Loop Networking.

In this methodology a single unidirectional line is run passing through each area that needs access to the network; access is achieved by nodes or breakout boxes of varying complexity depending on the requirement, each of which has a unique address for programmability purposes. At the simplest, a small fixed number of inputs to the network and outputs from the network may be offered at the node, along with unique control data. These may be analog ins/outs or digital ins/outs or a combination, and each may either look at (in the case of outputs) any of the (say) 64 program slots, or select a slot into which to place their audio (in the case of inputs).

More advanced nodes could, by way of example, look at many slots, mix them, mix that with local input material, and even place the composite mix into a slot(s) in the network. A frequent application is to retrieve audio from a slot and replace it with local input material. Signal processing specific to a local need (e.g., crossover/EQ for a speaker cluster) can be done within such a node; indeed, such products can be thought of primarily as a processor that just happens to have wide connectivity through the network and is marketed as such. Third-party or multiple vendors can be interoperable, providing they're all licensees of the same networking protocol.

This modified stream is then sent downstream to the next node, and so forth. The stream can be unidirectional (serial) or looped back upon itself (surprise, loop), whereupon the originating node seemingly perversely sees as its input the stream after it has passed through all the other nodes.

Such networks tend to be quite efficient, since they are able to reuse slots along the way. Disadvantages are:

- That serial anythings tend to be badly affected by single-point failures—in other words, one node failing makes orphans of all the others downstream of it, bisecting the net.
- It takes an appreciable time to receive the packet of slots, disassemble it, modify slots to whatever degree, reassemble it, and send it on its way, and such processing latency is obviously cumulative with that from previous and successive nodes. That said, latencies can be low, in the handful-of-sample-period range, trivial compared to those of 100 MHz TCP/IP systems.
- The network cabling routing has to be carefully thought through and follow a logical progression of where the audio needs to go next. This sometimes isn't easy.

25.25.3.3.2 Star Network Topology.

In contrast to relying on the packet addressability of an IP-style Ethernet network—which belies the need for centralized command and control—UDP-style networks which dumbly if faithfully and with low fixed latency propel a fixed amount of audio from one end to the other of a straight pipe, requires a central switch: this unpacks each incoming stream, decides which elements within them need to go where, and assembles outgoing streams appropriately. Such routing systems were long commonplace in broadcast installations, and the morphing of the concept to using high-bandwidth UDP pipes was a natural and welcome progression from running lots of discrete signal lines.

Having a central hub from and to which all network runs are connected is a concept as old as the phone system. As such, it has similar strengths and weaknesses in that cable runs are logical and obvious, but reliability hinges on that of a central server. Although this would seem a vulnerability, a single point of failure, parallel and redundant methodologies are common, as will be seen.

The switch in audio terms is in fact a router, which accepts large numbers of sources and can redistribute them in any combination to large numbers of destinations. Ofttimes many sources and destinations are local to the router, but more often high-density audio pipes as described above of, say, 64 discrete signals bidirectionally, spur out to remote locations, where these pipes are terminated in input and output terminations of whatever nature and complexity are desired. (If all the outputs need to be analog on XLRs, so be it. AES pairs, no problem—termination styles are easily accomplished to suit the application. If local-specific signal processing is desired, no problem.) These pipes are simply arranged to look like multiple sources and destinations to the router and are treated as such. The router also parses any metadata, logic control, or metering that accompanies each audio path and routes or deals with each of these accordingly. (Losing the meta data or sending it to the wrong place is like the airline losing your bags: it isn't the end of civilization, as you've arrived, but you're nowhere near as equipped to accomplish what you have to do.)

This centralized router model works well in the broadcast environment, radio or TV, where much of the engineering work is clustered in a central racks room anyway; likely many of the sources and destinations of the router would be local to it in that room, easing interconnection, with spokes of high-density audio transport issuing out to each studio and production area—Something of a natural for the star topology.

Live sound benefits from this method more than others (such as serial), too. A typical setup is for two consoles (house and monitors) to each be recipients of all stage audio sources or at least major subsets of them. In this instance the router would receive the outputs from the active stage boxes as sources and distribute them as required by transport links to the two consoles.

Returning from the consoles to the router are:

1. House—main mixes and/or speaker processor outputs
2. Monitors—many, stage monitor mixes

These enter the router and are then sent by further links to the desired amplifier racks for the flown and/or stacked PA and sub-low cabinets, or indeed straight to the powered speakers themselves; and to the amplifier racks for the stage monitor speakers and the transmitter rack for in-ear monitoring.

Additional feeds, such as for recording, are assembled in the router and sent to the band's DAW or the recording truck and terminated accordingly.

25.25.3.3.3 Mixing in the Router

A progression from the notion of a router being a simple crossbar or summing switch is that it becomes a soft matrix, where the relative levels of sources and destinations may be varied. In other words, a mixer. Or a number of smaller mixers than the total capability described by possible numbers of inputs and outputs.

And, taking a step further forward, given the already signal-processing-intensive environment, the console-style signal processing on mixer inputs, outputs, and submixes becomes relatively easy to implement.

Nodes in a TCP/IP system, or breakout boxes in serial network schemes, often contain varying amounts of processing, all the way from simple access to slots to being a full-scale mixer. It is common for the router to have processing capability, indeed for it to be where the mixing/processing parts of consoles reside; it makes sense, since all the possible component signals to be mixed exist within the router or can be got there expeditiously enough through one or several links. Fig. 25-155 shows a large-format TV audio mixing console in which there is not a shred of audio—it is merely a control surface, controlling the signal processing/mixing elsewhere within a processing router: Fig. 25-156. That the console system is or is part of a router means that the number of available mix sources is limited only by the size of the router and can so extend to the thousands. From an operating perspective, however, the console is limited to mixing only as many sources instantaneously as it has faders; this can be multiplied by paging the surface, such that it can flip-flop either entirely or on a fader-by-fader basis to control multiple other channels; instantaneous channel counts can thus run into the hundreds, and on big live shows (such as election night) often does.

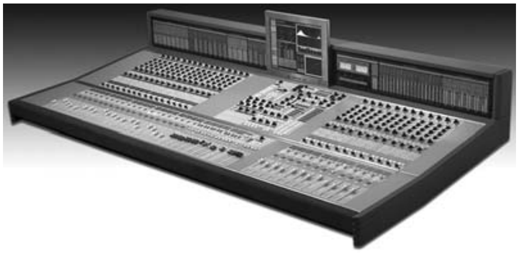


Figure 25-155. The Wheatstone D5.1 large-format TV audio console, a control surface which has no audio in it at all! It merely controls a remote router, the "Bridge" in Fig. 25-156. Courtesy Wheatstone Corporation.

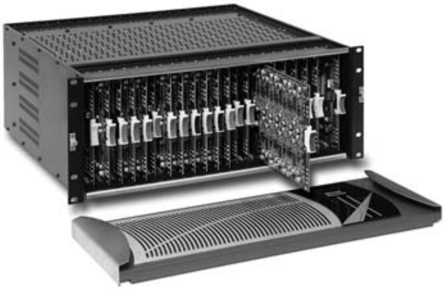


Figure 25-156. A Wheatstone "Bridge" cage, a mixing/processing router providing the audio "engine" for the D5.1 console in Fig. 25-155. Courtesy Wheatstone Corporation.

There are two other major applications for this router-as-console solution. One is radio studios—several, if relatively small, consoles within a single complex: they can all share the same hardware and resources of a single router, indeed may all be in the same box, yet to all intents and purposes be discrete operationally.

It is a natural live sound solution, too, where as described above there may indeed be multiple fair-sized console systems (house, monitors, recording) but that all share common sources from the stage and yet have very separate destinations (PA, monitors, recorder). The mixing router not only performs the signal routings, but it is also home to all the console-type signal processing required for all three operations.

A valid concern for each of the above applications is that of single point of failure. Large routing mixers typically address this with fail-safe measures, meaning each host microcomputer has a hot standby ready to take over if the main one should hiccup, and spare signal processing/mixing DSP boards equally stand ready to be reassigned on-the-fly to take over from one that may have halted. Some designs even have an entirely separate router, operating in parallel to the main one, ready to take over in the case of a failure. Although they could be perceived as expensive precautions, they look really inexpensive in relation to dead air.

Chapter 26

VU Meters and Devices

by Glen Ballou

26.1 General 997
26.2 Standard VU Meters 997
  26.2.1 Meter Ballistics 999
  26.2.2 Reference Levels 1000
  26.2.3 VU Meter Impedance 1000
  26.2.4 VU Impedance Level Correction. 1000
  26.2.5 Voltages at Various Impedances 1001
26.3 Wide-Range VU Meters 1002
26.4 Bar Graph VU and Spectrum Analyzers 1002
26.5 Power-Level Meters 1003
26.6 Power-Output Meters 1003
26.7 Peak Program Meters 1003
  26.7.1 DIN 45406 Standard 1003
  26.7.2 British Broadcast Standard 1004
  26.7.3 Nordic N9 Standard. 1004
26.8 AES/EBU Digital Peak Meter 1004
26.9 Loudness Meters 1005
26.10 Surround Sound Analyzer 1006
References 1008

This page intentionally left blank

26.1 General

To operate a sound recording or reproducing system properly, some method for determining the signal levels in different parts of the system to avoid overloading, noise, and distortion is required. This is the purpose of the *volume indicator (VI) meter*. A VI meter is a meter used to measure levels of audio-frequency signals. The term *volume indicator* is generally associated with meters calibrated in decibels. Until recently, volume unit (VU) meters were devices to measure power with respect to 1 mW of power across a 600 Ω line. Today VU measurements are made with respect to many different bases.

VU meters were first used by the telephone company. They were used to measure the level of the signal being sent down the line. The lines were open wire pair of AWG #6 wire spaced 12 inches apart, which translated to a characteristic impedance of 600 Ω as determined with the equation

$$Z = 276\log\left(\frac{2D}{d}\right) \tag{26-1}$$

where,

D is the spacing of the two wires,

d is the diameter of the wire.

Today, most amplifying devices have a high-impedance input and a low-impedance output as specified by a 1978 I.E.C. standard requiring the output impedance of a device to be less than 50 Ω and the input impedance to be greater than 10 kΩ. Since very little power is transferred between 50 Ω and 10 kΩ, it makes more sense to make measurements as voltage gain rather than power gain.

It is important to know what kind of measurement reference is being used. The following are some of the common references:

dBm. The original definition of the dB. It is power level in dB referenced to 0 dB or 1 mW and a 600 Ω load.

dBW. Power referenced to 1 watt.

dBf. Power referenced to 1 femtowatt (1×10^{-15} W).

dBV. Voltage referenced to 1 Vrms. dBV is not affected by impedance.

−10 dBV. A voltage reference level used by many consumer products and is equal to 0.316 Vrms.

dBu. Voltage referenced to 0.775 Vrms. It is not affected by impedance. The u stands for unterminated.

+4 dBu. The pro-audio voltage reference level of 1.23 Vrms.

dB FS. Digital audio reference level equal to full scale, which is the maximum peak voltage level before digital clipping of a data converter. Full-scale value varies with each design.

dBA. An unofficial method of stating loudness measurements using the "A" weighted curve on a sound level meter.

dBC. An unofficial method of stating loudness measurements using the "C" weighted curve on a sound level meter.

dB-SPL. Sound pressure level referenced to 0.0002 µbar where 1 µbar = 1 dyne/cm^2 or the threshold of hearing.

dBr. An arbitrary reference level that must be specified. It can be used for many different references as long as it is specified.

DIN Scale. The DIN scale as used in Germany and Austria uses +6 dBu as the reference level for the 0 dB mark. This is equivalent to 1.55 Vrms.

26.2 Standard VU Meters

A *volume unit (VU) meter* is a special form of VI meter used for monitoring broadcast, recording circuits and sound reinforcement systems. Such meters employ special ballistics that average out complex waveforms to properly indicate program material that varies simultaneously in both amplitude and frequency. For complex waveforms, such as speech, a VU meter reads between the average and the peak values of the complex wave. No simple relationship exists between volume measured in VU and the power of a complex waveform. The indicated reading will depend on the particular wave shape at the moment. For sine-wave measurements, a change of one VU is numerically equal to a change of 1 dB.

VU meters are designed to have a dynamic characteristic that approximates the response of the human ear. When a speech waveform is applied to a VU meter, the movement will indicate peaks and valleys in the signal. The average of the three highest peaks in 10 s (disregarding occasional extremes) is taken to be the indication of the meter movement.

Many meters marked as VU meters are not actually such meters, since they do not have the special ballistics and characteristics of the standard VU meter.

The *VU* meter is a device whose standard has remained the same since 1961. The meter consists of a 200 mAdc D'Arsonval movement fed from a full-wave, copper-oxide rectifier mounted within the meter case. VU meters are calibrated in reference to 1 mW of power into a 600 Ω load. A typical moving coil VU meter is shown in Fig. 26-1.

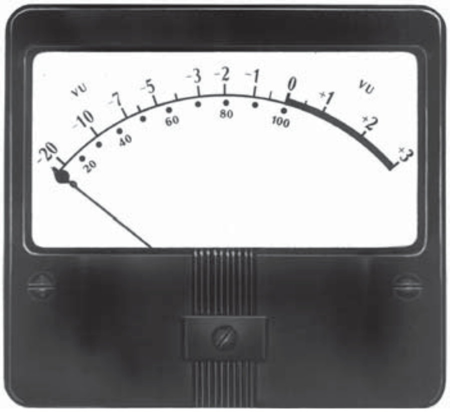
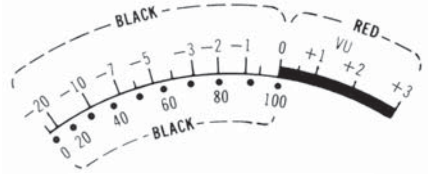


Figure 26-1. Moving coil VU meter. Courtesy Simpson Electric.

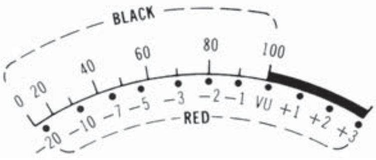
In the 1920s and 30s copper-oxide rectifier power-level meters were inaccurate and not satisfactory for program monitoring. The development of an entirely new meter was jointly undertaken by the Bell Telephone Laboratories, Columbia Broadcasting System (CBS), and the National Broadcasting Company (NBC). The results of this research were not only the development of a new type VI meter but also the standardization of a new reference level of 1 mW, a unit that was adopted by the electronics industry in May 1939. The current standard is ANSI C16.5-1961, formerly the Acoustical Society of America (ASA) C16.5-1961.

The characteristics of the dBm VU meter are as follows:

- **General.** The meter consists of a dc meter movement with a full-wave, copper-oxide rectifier unit (mounted in the instrument case) and responds approximately to the root-mean-square (rms) value of the impressed voltage. This value will vary somewhat depending on the waveforms and the percentage of harmonics present in the signal.
- **Instrument Scale.** The face of the instrument may have either of the two scale cards shown in Fig. 26-2. Each card has two scales: a VU scale ranging from −20 to +3 VU and a percent-modulation scale ranging from 0 to 100%, with 100% coinciding with the 0 point on the VU scale. The normal point for reading volume levels is at 0 VU or 100%, which are located to the right of the center at about 71% of the full-scale arc.



A. Recording and test equipment.



B. Broadcast monitoring.

Figure 26-2. VU meter scales.

- **Dynamic Characteristics.** With the instrument connected across a 600 Ω external resistance, the sudden application of a sine-wave voltage, sufficient to give a steady-state deflection at the 0 VU or 100 scale point, shall cause the pointer to overshoot not less than 1% or more than 1.5% (0.15 dB). The pointer shall reach 99 on the percent scale in 0.3 s.
- **Response Versus Frequency.** The instrument sensitivity shall not depart from that at 1 kHz by more than 0.2 dB between 35 Hz and 10 kHz, or more than 0.5 dB, between 25 Hz and 16 kHz.
- **Impedance.** For bridging across a line, the volume indicator, including the instrument and proper series resistor (3600 Ω), shall have an impedance of 7500 Ω when measured with a sinusoidal voltage sufficient to deflect the meter to 0 VU or the 100% scale point.
- **Sensitivity.** The application of a sinusoidal potential of 1.228 V (4 dB above 1 mW in a 600 Ω line) to the instrument in series with the proper resistance (3600 Ω) will cause a deflection to the 0 VU or 100% point.
- **Harmonic Distortion.** The harmonic distortion introduced in a 600 Ω circuit, caused by bridging the volume indicator across it, is less than 0.3%, under the worst possible condition (no loss in the variable attenuator).
- **Overload.** The instrument must be capable of withstanding, without injury or effect on the calibration, overload peaks of ten times the voltage equivalent to

a reading of 0 VU or 100% for 0.50 s and a continuous overload of five times that voltage.

26.2.1 Meter Ballistics

Meter ballistics are the mechanical and electrical characteristics built into the meter movement. A given characteristic may be obtained by shaping the pole pieces and counterweighting the pointer mechanism. Shunts are sometimes used across the meter terminals, but this use will reduce the sensitivity of the movement.

The ballistics characteristics of a typical old-style VI meter or voltmeter and a standard VU meter, when a 1000 Hz signal is applied for a period of 1 s, are shown in Fig. 26-3. Note the VU meter comes to a steady state at the end of 0.30 s, while the VI meter continues to oscillate showing peaks and valleys over a period of 1 s. An ac voltmeter would be even worse than the old style VI meter as it would never settle down and would constantly overshoot. This clearly indicates why the ballistics of the VU meter are desirable for monitoring program material containing complex waveforms.

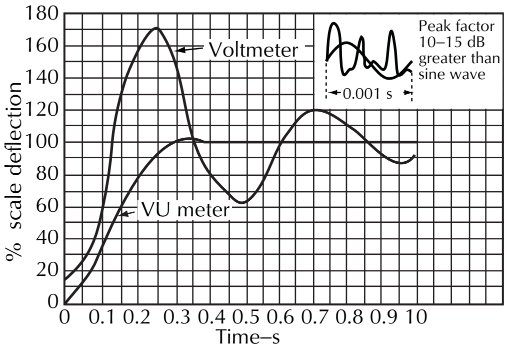


Figure 26-3. Comparison of the original VI meter and the present VU meter ballistics when a 1000 Hz signal is applied for 1 s.

A VU meter reads the rms value of the waveform. On a sine wave, the rms VU indicator of the peak is only 3 dB above the reading; however, on voice or music, the peak may be 10–12 dB above the VU reading. This difference is called the *crest factor* and is illustrated in Fig. 26-4.

Because of the meter ballistics, a VU meter indicates somewhere between the average and the peak values. Program material is of a complex and transient nature; therefore, the VU meter reading is considerably under the instantaneous peak program level. This means that 8–14 dB peaks present in the program material are not indicated by the meter because the meter movement cannot follow small instantaneous peaks. Even if they could be seen, it would be too late to reduce the level. Therefore, the meter must either be set or caused to indicate in a manner that will not overload the system in which it is operating.

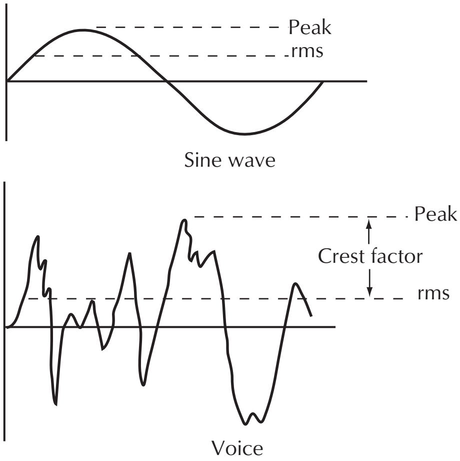


Figure 26-4. Crest factor caused by the peak of music or voice being greater than $\sqrt{3}$ rms.

Since VU meters do not include the true peak values of program material (complex waveforms), it is quite easy to overload a recording system. To protect against these unseen peaks, a lead or margin of safety is inserted in the VU meter circuit.

To insert a lead into a VU meter circuit, the VU meter is connected across a bridging bus with a sine-wave level of +14 dBm. A 400 Hz or 1000 Hz signal is sent into the input of the recording console. The mixer control is set to its normal operating range, and the signal level is adjusted to bring the bus level to +14 dBm (the VU meter reads 100% or 0 dBm).

Remove the input signal and return the VU meter attenuator to its +6 dBm position. This inserts an 8 dB lead or margin of safety in the VU meter by making it 8 dB more sensitive. Thus, it protects the system against unseen peaks up to 8 dB. The program material is now mixed in the usual manner. Some recording activities, because of the heavy peaks and overloads encountered in some types of music, use a 10–12 dB lead in the VU meter.

Radio transmitters are adjusted in a similar manner. However, in this instance, the percent modulation indicated by the VU meter indicates the percent modulation of the radio transmitter.

26.2.2 Reference Levels

In the early days of broadcasting and recording, both 10 mW and 12.5 mW into a 500 Ω line were used as a reference level. However, later this was changed to 6 mW. In May 1939 the present standard of 1 mW into a 600 Ω line was adopted. This reference level was selected as a level that would conform to the telephone company's standards of limiting the signal level on a transmission line to a value that would produce a minimum of crosstalk and still provide a satisfactory signal-to-noise ratio (SNR). The 1 mW reference level is a unit quantity and is readily applicable to the decimal system, being related to the watt by the factor 10^{-3}.

Zero level is a reference power level of 1 mW of power into a 600 Ω load. This is equivalent to a voltage of 0.775 V.

26.2.3 VU Meter Impedance

The VU meter and its attenuator impress a 7500 Ω impedance onto a circuit. The VU meter system consists of an indicator movement, a variable attenuator, and a series resistor of 3600 Ω, Fig. 26-5. Meter manufacturers supply only the meter movement; the external circuitry is added later. A 200 μA D'Arsonval meter movement with an internal resistance of 3900 Ω and a full-wave, copper-oxide or selenium rectifier are contained within the meter case. The attenuator is variable in steps of 2 dB, presents a constant impedance of 3900 Ω to the meter movement, and prevents the ballistics of the meter from being affected when the attenuator setting is changed.

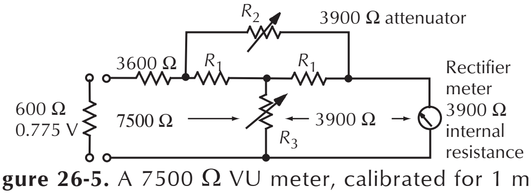


Figure 26-5. A 7500 Ω VU meter, calibrated for 1 mW reference level or 0.775 V across 600 Ω.

Standard VU meters are designed to read 0 VU, or 100%, with 1.228 V (+4 dBm) applied to the instrument. If the meter is used with the attenuator but without the 3600 Ω series resistor and is connected across a 600 Ω load in which 1 mW of power is flowing, the movement will be deflected to the 100% calibration point. This method is not recommended since the impedance looking back into the meter is only 3900 Ω and loads the 600 Ω circuit. It is the usual practice to keep the impedance of bridging devices at a ratio of 10:1 or greater.

Increasing the input impedance of the VU meter from 3900 Ω to 7500 Ω creates a 4 dB loss across the 3600 Ω resistor. If a signal of 1 mW (0.775 V) is impressed across the input terminals of the circuit in Fig. 26-6, it will not deflect the meter to the 0 VU calibration but only to the −4 VU (or decibel) mark, or approximately 65%. This means that if the meter is to be deflected to the 100% point, the input signal must be increased to a +4 dBm. This is the reason why 1 mW of power will be indicated at the −4 dB calibration mark.

Attenuators used with VU meters start at a +4 dBm. The bridging loss caused by the VU meter being inserted into the circuit is the drop in signal level caused by the absorption of power by the meter circuit. As a rule, the power absorbed is quite small and may be ignored. However, at high powers, it may become important. Bridging loss may be calculated by the equation

$$dB_{loss} = 20\log\frac{2B_R + Z}{2B_R} \tag{26-2}$$

where,
B_R is the VU meter input impedance,
Z is the line impedance.

A 7500 Ω VU meter has a bridging loss of 0.34 dB.

26.2.4 VU Impedance Level Correction

VU meters are calibrated for 1 mW of power across a 600 Ω load as −4 VU, therefore when a VU meter is connected across any other impedance, a correction must be added to the indicated reading to give a proper VU reading. The equation for the level correction is

$$dB_{corr} = 10\log\frac{Z_2}{Z_1} \tag{26-3}$$

where,
dBcorr is the decibel amount added to the VU reading,
Z_2 is the impedance for which the meter is calibrated,
Z_1 is the impedance of the circuit bridged.

A typical example of applying a correction factor is as follows: a VU meter calibrated for a line impedance of 600 Ω is bridged across a 16 Ω loudspeaker line and indicates a level of +1 dBm. The true VU would be

$$VU = 1\text{ dBm} + correction\ factor\,. \tag{26-4}$$

The correction factor from Eq. 26-2 is

$$
\begin{aligned}
dB_{corr} &= 10\log\frac{600}{16}\\
&= 10\times 1.574\\
&= 15.74\text{ dB}
\end{aligned}
$$

The correction factor of 15.74 dB is added to the meter reading of +1 dBm for a true level reading of +16.74 dBm. Typical correction factors are shown in Table 26-1.

Table 26-1. Correction Factors in dBm to Be Applied to a VU Meter When Connected Across an Impedance Other Than 600 Ω

Line Impedance—Ω	Meter Cal 600 Ω—dB
10,000	−12.22
5000	−9.21
2500	−6.20
1000	−2.22
600	0.000
500	+0.791
250	+3.800
200	+4.770
150	+6.020
125	+6.810
100	+7.780
50	+10.790
30	+13.010
16	+15.740
15	+16.020
8	+18.750
4	+21.760

If a VU meter is connected across a line impedance different from that for which it was originally calibrated, the voltage supplied to the meter will either be lower or higher than the original calibration; therefore, the meter would indicate incorrectly. Two circuits are shown in Fig. 26-6, one a 600 Ω circuit and the other a 16 Ω circuit. Both are dissipating the same amount of power; yet the voltage for the 600 Ω circuit is 0.775 V, and for the 16 Ω circuit it is 0.127 V. As can be seen, if a VU meter is connected across the 16 Ω circuit, it will not deflect the same amount as for the 600 Ω circuit, although the same amount of power is flowing in each circuit. To arrive at the correct power level in the 16 Ω circuit, a correction factor must be applied to the meter indication.

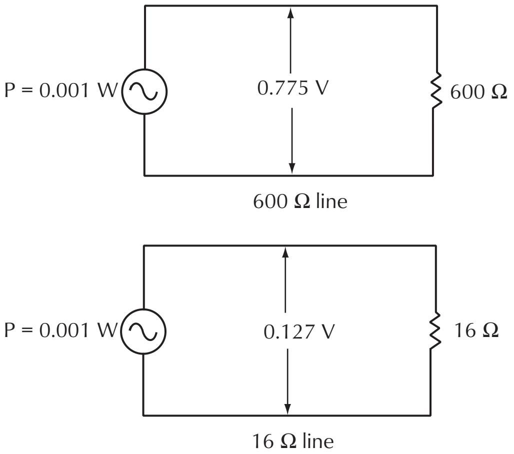


Figure 26-6. Voltage across lines of different impedance but with the same power in milliwatts.

26.2.5 Voltages at Various Impedances

If the line voltage for a given level at 600 Ω is known, voltages for other line impedances may be calculated using

$$V_x = V\sqrt{\frac{Z}{600}} \tag{26-5}$$

where,

V_x is the unknown voltage,

V is the voltage for 600 Ω,

Z is the new impedance.

As an example, assume voltage V_x is required for a line impedance of 150 Ω at a level of +4 dBm. Referring to Fig. 26-7, the voltage for a level of +4 dBm at 600 Ω is 1.23 V. The new voltage may now be calculated using

$$
\begin{aligned}
V_x &= 123\sqrt{\frac{150}{600}}\\
&= 0.615V
\end{aligned}
$$

Voltages for a line impedance of 600 Ω for levels between 0 and +50 dBm may be taken from Fig. 26-7. Voltage across 600 Ω can be calculated from dBm with the following equation

$$V = 0.6\times 10^{\frac{dBm}{10}} \tag{26-6}$$

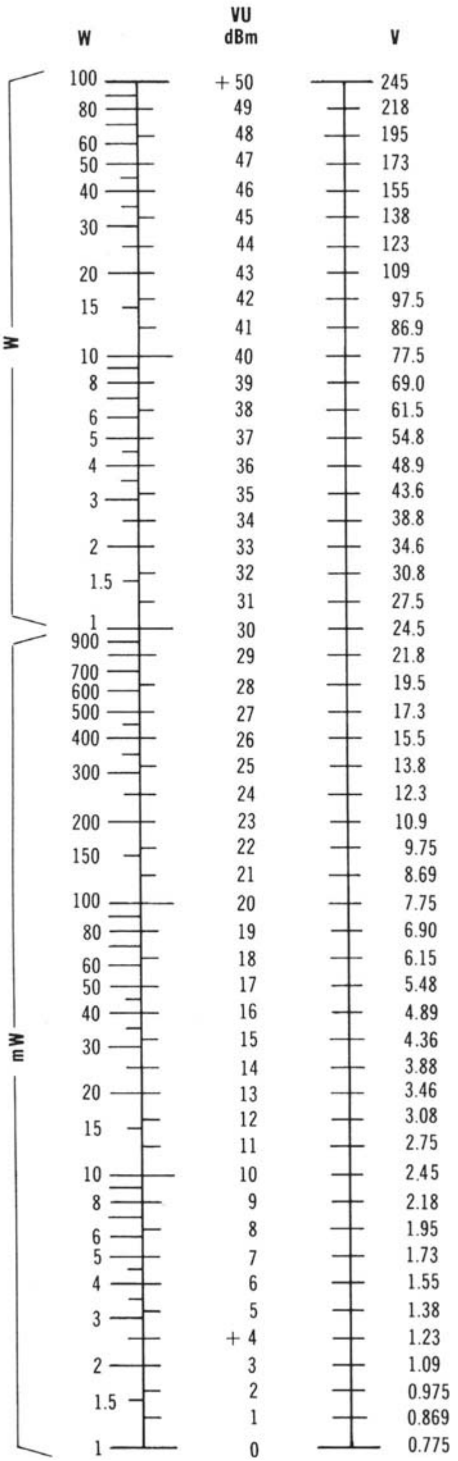


Figure 26-7. Relationship of VU and dBm to power in watts and voltage in a 600 Ω line.

26.3 Wide-Range VU Meters

Standard VU meters measure only the upper 23 dB of the signal level. From the practical standpoint, this limits the display to about 20 dB below the reference level of the 0 indication.

This short range of operation limits its usefulness, particularly when it is connected across a bridging bus for monitoring program information. A *wide-range program-monitor meter*, Fig. 26-8, displays the program information over a 60 dB meter scale, spread from –57 dB to +3 dB. The large spread of program material permits the very low-level signals to be observed as well as the noise between program pauses. The wide-range VU meter was not designed to replace the conventional VU meter; however, its characteristics are compatible with the VU meter. In addition, a dc output is provided for connection to a linear tape recorder for logging program levels over a range of 60 dB. The 0 dB indication may be set to represent a reference level from –22 dBm to +18 dBm.

The basic component is a logarithmic amplifier, Fig. 26-9, with a nonlinear feedback circuit, a preamplifier, a 15 kΩ bridging input transformer, a reference-level selector switch, and a sensitive indicating meter movement.

26.4 Bar Graph VU and Spectrum Analyzers

The United Recording Electronics Industries (UREI) Model 970 Vidigraf is a bar graph display generator that operates any National Television System Committee (NTSC) standard video monitor or (with an inexpensive accessory) black-and-white television receiver. The system provides both a VU level display and the frequency-spectrum-level information. It is designed primarily for multitrack recording studio applications. However, its dc to 20 kHz input capability suggests its use for a wide range of dc or ac analog voltage measurements.

The 970 Vidigraf's modular construction provides users with complete flexibility to adapt the system to their specific needs. A maximum of four 16-channel input display modules may be installed for VU level, automation control voltages, or frequency-spectrum viewing. Each module may be individually switched to the video generator in the single mode. In the dual display mode, the screen is split vertically to accommodate the information from any two input modules simultaneously. Instantaneous identification of the input channel sources and/or frequencies, as well as vertical scaling indices are automatically provided by the built-in

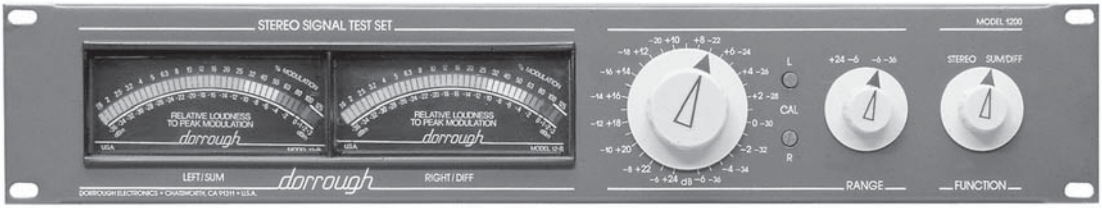


Figure 26-8. Wide range VU meter. Courtesy Dorrough Electronics.

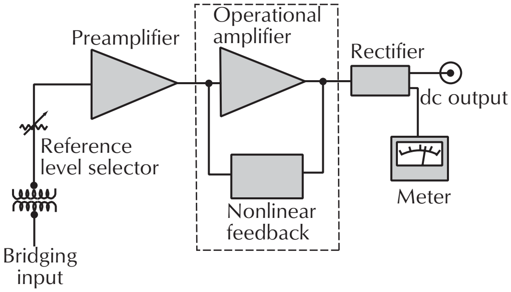


Figure 26-9. Block diagram of a wide-range program monitor VU meter.

programmable character generators. This eliminates any need for screen overlays or masks and ensures accurate positioning of the alphanumeric information regardless of screen size or width and height adjustments.

Some typical displays are:

- 6 or 32 simultaneous VU channels.
- 16 or 2 × 16 bands of frequency spectrum (1 or 2 channels).
- 16 VU channels, plus channels of automation control voltages.
- 16 VU channels, plus 15 bands of frequency spectrum and 1 composite level.

One *VU module* provides 16 bar graphs with standard VU ballistics over a display range of 30 dB. Each bar has two shades of gray, with the lighter shade above the 0 dB reference. When a signal is applied to any of the 16 inputs, a bright bar moves up and down with the signal level. The 0 dB reference point can be continuously adjusted to any standard from 0 to +8 dB. The VU module is user programmable to display a logarithmic scale of –20 dB to +3 dB when measuring audio signals or to read linearly from 0 to 10 for display of ac or automation dc control voltages.

The *spectrum module* provides a visual real-time display of VU level versus frequency of an audio signal. It is useful for setting equalization and adjusting frequency balance. This module provides 16 bar graphs with visual characteristics similar to those of the VU module. One bar is assigned to the full spectrum of the audio signal, and the other 15 channels display increments of the frequency spectrum, centered on standard ISO ⅔-octave filter frequencies. Two independent controls adjust the level of the full spectrum bar relative to the spectrum analysis bars.

26.5 Power-Level Meters

A *power-level meter* is a VI meter calibrated in decibels. As a rule, this type of meter is normally used with test equipment for steady-state measurements and is not used for monitoring program material because its ballistics are more like those of a voltmeter.

26.6 Power-Output Meters

A *power-output meter* is used for measuring the power output of audio amplifiers and other devices. It may also be used to determine the characteristic and internal output impedance, the effect of load-impedance variation, and other applications involving the measurement of output power and impedance with respect to frequency. The power output meter may be calibrated in watts and/or dBm. The power output meter is a test instrument and not used for monitoring program level because of its ballistics.

26.7 Peak Program Meters

The *peak program meter* (PPM) is used extensively in Europe and falls under four standards, the DIN-type DIN 45406, the BBC type, the EBU type, and the Nordic N9 type. These meters measure the peak program signal, which is usually +6 dB to +20 dB above the readings seen on the VU meter.

26.7.1 DIN 45406 Standard

The PPM is popular in Europe. It is designed to have a fast rise time, 30 times as fast as a VU meter, and a much slower fallback or decay time.

The DIN 45406 and the IEC 268-10 have an integration time of 10 ms and a decay time of 1.5 s for 20 dB of fallback and 2.5 s for 40 dB of decay. The indicator range is –50 dB to +5 dB. The scale marked for 100% reading is 0 dB which is the reference level of +6 dBu or 1.55 Vrms, Fig. 26-10A.

The RTW 1019GL analog peak program meter + loudness meter + phase correlation meter is shown in Fig. 26-10A. The 127 mm (5 inch) 201 element bar graph display has a 1 dB per division scale from +5 dB to –10 dB changing logarithmically to –50 dB. Roll-off above 20 kHz is 12 dB/octave. The meter includes a +20 dB gain increase and a peak memory/reset circuit. The integration time is selectable between 1 ms and 10 ms. The balanced input is transformer isolated.

The meter panel also includes a three-color phase correlation value display with memory. The correlator indicates the phase correlation r, of stereo signals. If both channels are in phase—e.g., a mono signal on both channels—the reading is +1 r. With only one or no signal at the input, the meter will read 0 r.

26.7.2 British Broadcast Standard

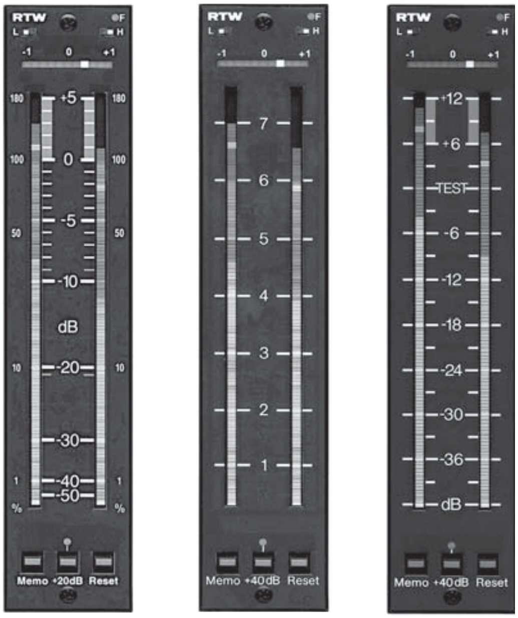
The British Broadcast Standard, BS 55428 Part 9, has an integration time of 12 ms and a decay time of 2.8 s for decay from 7 to 1. The indicator range of 1 s to 7 is equivalent to a –12 dB to +12 dB. The scale mark for 100% reading is 6 and is referenced to +8 dBu or 1.95 Vrms.

Fig. 26-10B shows an analog RTW 1034GL British standard scale IIa analog peak program meter + loudness meter + phase correlation meter. The 127 mm (5 inch) 201 element bar graph display measures from –12 dB to +12 dB. The meter includes a +40 dB gain increase and a peak memory/reset circuit. The integration time is selectable between 1 ms for digital audio and 10 ms for analog audio. The balanced input is transformer isolated. The meter panel also includes a three color-phase correlation value display with memory.

26.7.3 Nordic N9 Standard

The Nordic Recommendation N9 has an integration time of 5 ms, a decay time of 1.7 s for 20 dB, and 3.4 s for 40 dB of decay. The indicator range is from –42 dB to +12 dB. The scale mark for 100% reading is 0 dB and is referenced to +6 dBu or 1.5 Vrms, Fig. 26-10C.

Fig. 26-10C shows an analog RTW 1039GL Nordic Recommendation N9 analog peak program meter + loudness meter + phase correlation meter. The 127 mm (5 inch) 201 element bar graph display measures from –42 dB to +12 dB. The meter includes a +40 dB gain increase and a peak memory/reset circuit. The integration time is selectable between 1 ms for digital audio and 5 ms for analog audio. The balanced input is transformer isolated. The meter panel also includes a three-color phase correlation value display with memory.



A. Stereo DIN 45406 peak program meter plus loudness meter plus phase correlation meter.

B. British standard stereo peak program meter plus loudness meter plus phase correlation meter.

C. Nordic N9 stereo peak program meter plus loudness meter plus phase correlation meter

Figure 26-10. European VI standards. Courtesy RTW GmbH & Co.KG, Cologne.

26.8 AES/EBU Digital Peak Meter

With the advent of digital equipment, new meter standards are being written to work with the AES/EBU digital format. This requires being capable of sampling 32 kHz, 44.056 kHz, 44.1 kHz, 48 kHz, and 96 kHz with an AES/EBU digital format.

The attack time is one sampling period and the decay time is 1.5 s for a change from 0 dB to –20 dB. The indicator range is from 0 dB to –60 dB.

The RTW 11529G digital peak program meter + loudness meter + phase correlation meter, Fig. 26-11A, has a 127 mm (5 inch) 201 element bargraph display. Its sampling rates are from 27 kHz to 96 kHz and it includes a dc filter and has indicators for 44.1 kHz, 48 kHz, and 96 kHz, emphasis, error, and overload. The meter also includes peak memory, peak hold, +40 dB gain, and a three-color correlation correction value display.

The RTW 11528G AES/EBU Digital PPM, Fig. 26-11B is especially useful for radio and TV broadcasting applications. The meter features AES/EBU inputs and outputs. The digital signal can be displayed once as it is without any weighting (sample precise display), which corresponds to the digital standard that has a scale range from –60 dB to +9 dB but with a fixed head room of –9 dB FS, which is marked 0 dB, and highlighted and superimposed with an integration time of 10 ms. It can also be displayed with a superimposed and highlighted loudness display. Finally, it can be shown as a 10 ms integration time–only function, as a quasi analog display. Its sampling rates are from 27 kHz to 96 kHz.

Both the RTW 11529G and the 11528G include an over indication with a selectable overload detector range, 9 to 24 bit overload response word length, and number of overload samples.

26.9 Loudness Meters

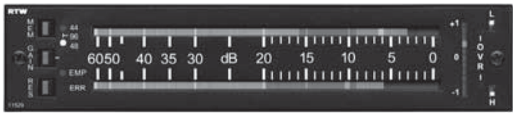
Loudness meters place VU and PPM meters on a single panel, providing an indication of the entire dynamic condition of the signal. It also eliminates the condition that eyeball wobble could develop in the attempt to follow two adjacent meters with differing ballistics. The use of two pointers with such differing ballistics on a single scale would demonstrate that the PPM would read consistently higher levels than the VU meter, and the large differential of decay with respect to rise time of the PPM in comparison to the equal rise and decay times of the VU meter would also be difficult to interpret.

Three types of scales are used on loudness meters:

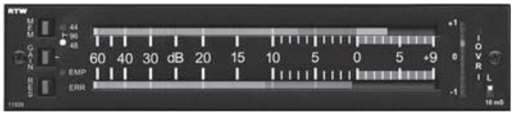
- Based on +14 dB of headroom.
- Referenced at 100% for broadcast transmission.
- Based on 20 dB of headroom.

The head room available to mixers in postproduction is not the same as allowed in broadcast. The U.S. standard in digital (SMPTE) is 20 dB below FS (full scale) and the EBU standard used in European and many Middle Eastern countries is −18 dB below FS. When film and post material is sent to the broadcast facility, the peak shall not exceed +12 dB analog or −8 dB digital.

The music and recording industries do not have these requirements for their products, and therefore use the full dynamic range. Commonly, this material will peak fairly consistently at −1 dB on a digital reading meter, with the bar graph fairly consistently four or five LEDs under the peak. If this material makes its way to broadcast, it will be louder than audio on video by possibly as much as 8 dB. The result might be rejection and need to redo the material using the guidelines required, or quality control at the broadcast facility will make a judgement on the loudness and turn it down accordingly. *Dialnorm* on HDTV was designed for these irregularities.



A. RTW 11529G Digital Peak Program Meter + Loudness Meter + Phase Correlation Meter.



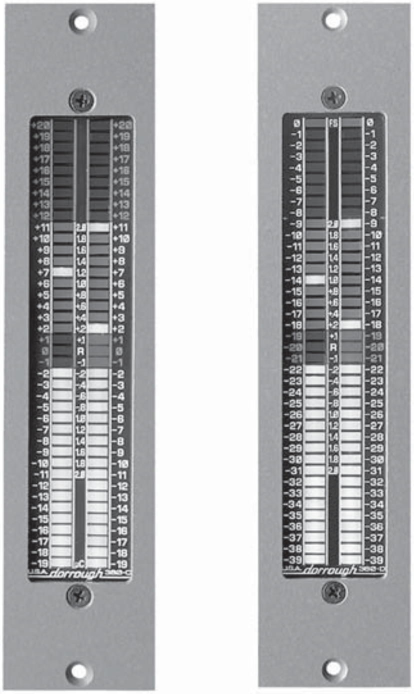
B. RTW 11528G Digital Peak Program Meter + Loudness Meter + Phase Correlation Meter.

Figure 26-11. An AES/EBU peak program meter plus loudness meter plus phase correlation meter. Courtesy RTW GmbH & Co.KG, Cologne.

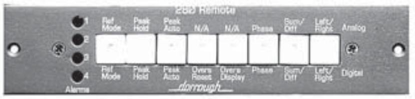
Observation of complex audio signals with an oscilloscope indicates the peak excursions of the program material. The use of an oscilloscope with a long or variable persistence CRT will show additional information relative to recurrent amplitude displayed by the persistence of the screen as a concentrated band of energy about the center of the CRT. It is these two pieces of information that provide the composition of acoustically related peak to quasi-average information. The meters shown in Fig. 26-11 feature a loudness display and peak holding display on each bargraph.

The meter in Fig. 26-12A is an analog reading meter, a digital reading meter is shown in Fig. 26-12B, and a remote control unit to access features from both is shown in Fig. 26-12C. The remote pushbuttons control the following functions:

- Left/Right.
- Sum/Difference.
- Phase.
- Overs Display with Overs Reset.
- Three second Peak Hold.
- Peak Hold Permanent.
- Reference Mode.



A. Analog reading B. Digital reading



C. Remote control

Figure 26-12. .The meter scale for a gain-riding device with both a normal persistence range (much like VU readings) and a normal peak range. Courtesy Dorrough Electronics.

The alarms on the left side of the remote are for Phase Error, Bit Stream Corruption, and Full Scale. Fig. 26-13 is a block diagram of the loudness meter of Fig. 26-12.

With signal input both meters read the audio in the same dynamic way. Each meter displays 20 dB of Peak Amplitude above the 0 VU persistence reference level, and therefore, 0 reference is the same on both.

The use of a dot display for Peak information and a bar display for Persistence information allows a single display for both ballistics. Each lamp in the display is therefore driven by two drivers, one for peak, the other for persistence. This representation presents a display of a dot riding on top of a bar graph. In order to make this display useful, there has to be a meaningful relationship between the two ballistics. The peak display has a rise time of two time constants, or 10 μs, which is 1000 times faster than the PPM. The decay time for the peak display is 18 ms/dB.

Equal energy, properly weighted for program material, will be discerned as equal in loudness. Since energy can be displayed as a function of amplitude and time, an oscilloscope can be used to confirm that large amplitudes of short tone bursts can be equal to longer tone bursts of lower amplitude.

Average power is defined as equal to the area under the curve divided by the time interval. Since the area is equal to the input energy, W or watts, during the interval,

$$P_{ave} = \frac{W}{t} \tag{26-7}$$

where,
W is in watts,
t is in seconds.

Thus, the averaging-type metering can provide an indication of power.

The persistence display has a time constant of 270 ms and a rise time of approximately 600 ms or twice as long as that of a VU meter.

Perceived loudness to the ear from source to source is determined by which circuit (peak or persistence) is first to illuminate its respective set of red LEDs. Program adjustments for equally perceived loudness should be holding either the peak or persistence excursions to its corresponding red LED area.

The relative loudness characteristic of the Peak to the bar graph has been retained by way of red LED reference points on both meters. This is the window of 12 dB of separation of the Peak from the Average for maintaining equal loudness. The +12 dB analog and –8 dB digital are the same scale points on both meters.

26.10 Surround Sound Analyzer

The surround sound analyzer method translates the important details of surround signals into a graphical display suited for instant evaluation. Successful mixing of surround signals is important. Besides the artistic and aesthetical aspects, there are fundamental technical preconditions for obtaining professional results.

Reality is often far from the ideal, particularly during live broadcasts and in audio production for video or TV. This makes it even more important to know, even in the most hectic working environment, how the surround mix will be perceived by the listener.

The RTW Surround Sound Analyzer—e. g., integrated in the RTW SurroundMonitor, Fig. 26-14—is a

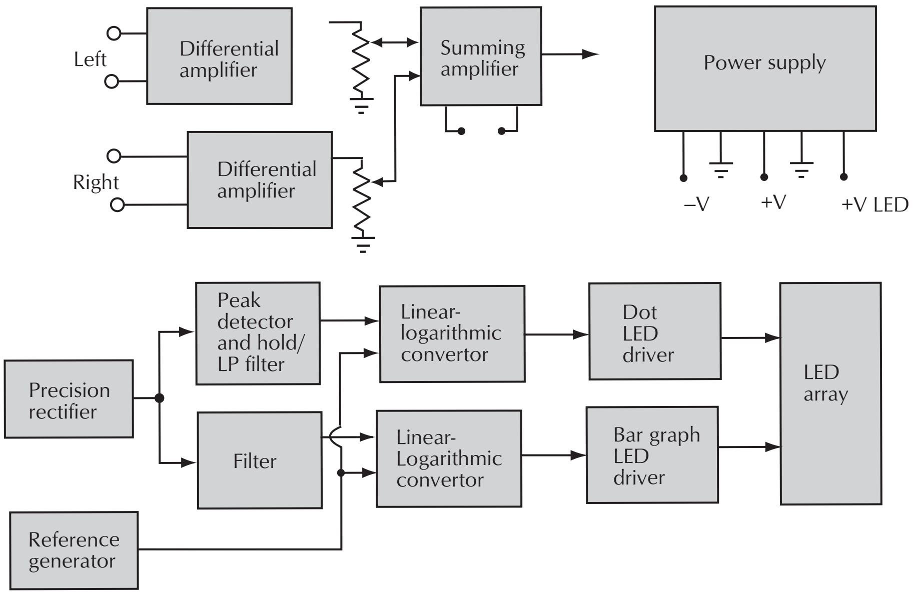


Figure 26-13. Block diagram of a Dorrough loudness monitor. Courtesy Dorrough Electronics.

unique tool showing all the important parameters of a surround signal at a glance. It gives detailed information for all individual channels as well as the overall effect of the mix.

The visual display of the Surround Sound Analyzer provides level and phase relations of all channels. The dynamic response of the display elements is a direct representation of the acoustic image so the balance of the surround sound can be observed.

The volumes of the four channels L, R, sL, and sR are displayed as diagonal white level bars originating from a common center point. Their tips are connected through cyan lines. The square formed by this figure, the total volume indicator (TVI), is a direct measure of the total volume and the balance of the acoustic image.

The curvature of these lines shows the channel correlation, positive values through an outward deflection (roof), negative values through an inward deflection (funnel).

The volume of the Center channel is indicated by another upwards-pointing level bar with yellow connecting lines, showing the perceptibility and dominance of the Center in relation to L and R.

Direction and width of front, side, and rear phantom sources are represented by lines between the loudspeaker symbols, called the Phantom Source Indicators (PSI). Their color changes with channel correlation. A separate correlation indicator for the two surround channels is available at the bottom of the display.

A cross representing the dominance vector indicates the position of the subjectively perceived center of gravity of the mix.

The Surround Sound Analyzer displays a correctly-scaled graphical representation of the relative volumes in the surround sound field. The interaction of levels (volume or sound pressure level) and the correlation of all channels in the production of the overall surround sound are displayed graphically.

The display in the Surround Sound Analyzer can be set to correspond to the volume or the reference sound pressure level by calibrating the instrument and the studio monitoring equipment accordingly. The axes of the 45° coordinate system use a dB volume level or dB-SPL scale and have a reference mark that is also displayed in the volume level and SPL displays in the peak program meter of the instrument. The balance between the Center channel and the L and R front channels is critical for all surround sound productions. The Center channel is displayed with its own display elements to show the volume differences between the Center channel and the L and R channels.

In addition to the signal level, the correlation level (and aligned to that the generation and the location of phantom sound sources) is important for multichannel sound productions, primarily in relation to downmixes or possible sound-faking erasements when generating a mono signal. The correlation level of the sL and sR surround channels is also important. Highly anomalous frequency-dependent correlation levels induce an unimpressive envelopment effect of the sL and sR surround signals. For monitoring this, the Surround Sound Analyzer features a correlation meter for the sL and sR surround channels. Fig. 26-15 shows the display and examples of various patterns.

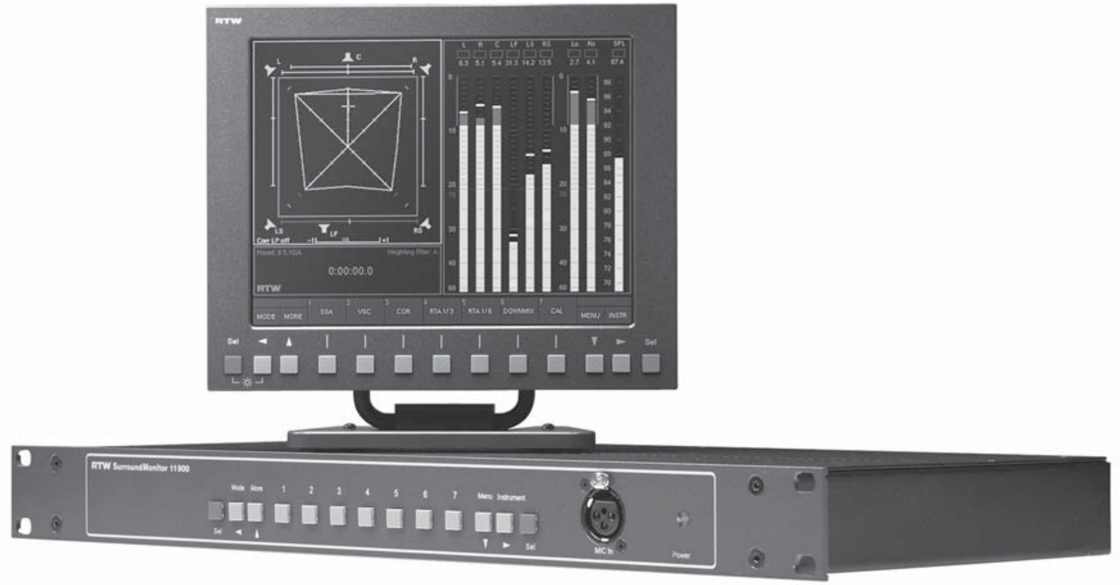

RTW
0:00:00.0
MODE
MORE
SSA
VSC
COR
RTA 1/3
RTA 1/6
DOWNMIX
CAL
MENU
INSTR
Sel
RTW SurroundMonitor 11900
Mode
More
1
2
3
4
5
6
7
Menu
Instrument
Sel
MIC In
Power

Figure 26-14. RTW SurroundMonitor 11900 for AES/EBU digital and analog standards with Surround Sound Analyzer. Courtesy RTW GmbH & Co.KG, Cologne.

References

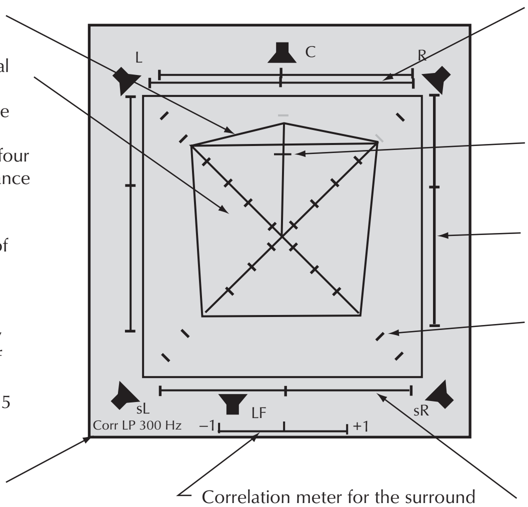
H. M. Tremaine, *Audio Cyclopedia*, Indianapolis, IN: Howard W. Sams & Co., Inc., 1969.

“VU Meter and Peak Program Meter—Peaceful Coexistence,” *J. Audio Eng. Soc.,* Preprint 1474 (G-5), Stephen F. Temmer, Gotham Audio Corporation.

“The Peak Program Meter and the VU Meter in Broadcasting,” *J. Audio Eng. Soc.*, Preprint 1691 (I-8), Hans Schmid, ABC.

“A Different Approach to the Old Problem of Audio Level Monitoring,” *J. Audio Eng. Soc.*, Preprint 2567 (B-1), R. W. Burden Associates, Canoga Park, CA, and M. Dorrough Electronics, Woodland Hills, CA, March 1988.

“The VU Meter,” *J. Audio Eng. Soc.*, Preprint 1794 (J-3), Michael Dorrough, Dorrough Electronics.



Center channel level indicator (yellow line)

Total Program Volume (TVI - Total Volume Indicator). The area enclosed by the lines is a measure for the total volume level, the spreading of the area across the four quadrants is an image of the balance of the sound characteristics. The shapes of the lines show the correlation level: a distinctive roof for r towards + 1 a funnel for r towards −1, a straight line for r between −0.25 and +0.25. The correlation level also is shown by the different colors of the lines of the phantom source indicators: green (+1 to +0.25), yellow (+0.25 to −0.25) and red (−0.25 to −1).

Switchable low pass filter for the sL – sR correlator

Correlation meter for the surround channels sL and sR

Moving bar indicators (PSI - Phantom Source Indicator) showing the position and width of the phantom sources between L and R and between C–L and C–R

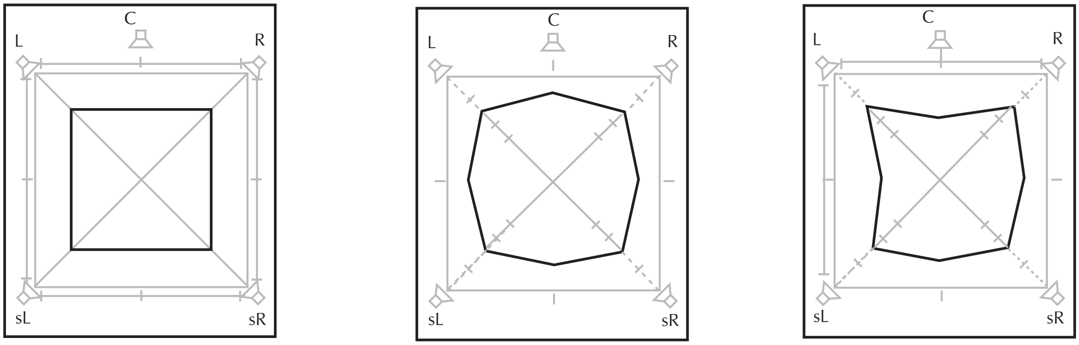
Indicator showing the position of the dominant sound event (dominance vector)

Moving bar indicators showing the position and width of the sidewise phantom sources (PSI)

Scaled coordinate system for the sound pressure level. The red mark refers to the reference monitoring sound pressure level like e. g. 78 dB(A) selected in the peakmeter section and to which the studio monitoring system can be calibrated.

Moving bar indicators showing the position and width of phantom surround sources (PSI)

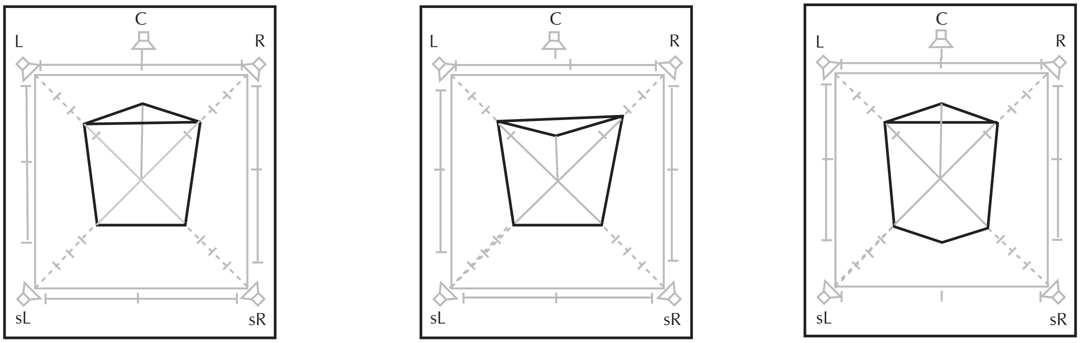
Actual screens are in color.



A. Incoherent noise, same level in the L, R, sL, and sR channels.

B. Sine wave signal, same level in the channels L, R, sL, and sR, similar to a mono signal.

C. Same as B but with the phase of the left channel rotated through 180°.



D. A surround signal with some Center presence.

E. A surround signal with a low level of Center presence.

F. The surround signal sL and sR is mono.

Figure 26-15. Surround Sound Analyzer screen views. Courtesy RTW GmbH & Co.KG, Cologne.

This page intentionally left blank

Part 5

Recording and Playback

This page intentionally left blank

Chapter 27

Analog Disk Playback

by George Alexandrovich and Glen Ballou

27.1 Introduction 1015
27.2 Disc/Record Dimensions 1015
  27.2.1 Record Diameter 1015
  27.2.2 Maximum Outer Diameter 1015
  27.2.3 Groove Dimensions 1015
  27.2.4 Stereophonic Groove 1015
  27.2.5 Channel Orientation 1016
  27.2.6 Channel Phasing 1016
  27.2.7 Channel Levels 1016
  27.2.8 Speed of Rotation 1016
  27.2.9 Lead-In Groove Pitch 1016
  27.2.10 Lead-Out Groove. 1016
  27.2.11 Finishing Groove. 1016
27.3 Signal Equalization in Disc Recording 1016
27.4 Turntables 1018
  27.4.1 Drive Systems 1018
  27.4.2 Turntable Design in the 21st Century 1021
27.5 Tonearms 1021
27.6 Phono Pickup/Transducers and Styli 1026
  27.6.1 Cartridge Styli 1031
  27.6.2 Cartridge Voltage Output 1033
  27.6.3 Electrical Loading 1033
27.7 Phono Preamplifiers 1033
27.8 Laser Turntable System 1034
27.9 Record Care Suggestions 1036
  27.9.1 Brushes 1036
  27.9.2 Record-Cleaning Machines 1036
  27.9.3 Record Storage 1037
  27.9.4 Cleaning Records 1037

This page intentionally left blank

27.1 Introduction

In the past 100 years approximately 30 billion phonograph records have been produced and sold. Music of the most famous composers and performers, orchestras, and bands, and sounds of events have been immortalized in intricate excursions of the analog record groove. Millions and perhaps billions of discs are still in the hand of the audiophiles, archives, musical libraries, DJs and radio stations.

The contents of all of these records can never be completely rerecorded onto the compact discs or another medium, so it is important that we can preserve, restore, and reproduce analog recordings.

The information contained in this section is directed toward the new generation of engineers and technicians so they may understand the reproduction techniques that led to digital technology. As we witness the decline in popularity of analog LP discs, remember that many developing countries around the world are still very much dependent on analog technology and in some cases what we consider the old 78 rpm format is the only source of prerecorded music and entertainment available to them.

Early recorded sounds had a high-frequency cutoff of 2–3 kHz. It took over 100 years to reach the sophistication of today's recording technology only to take a couple of steps backward in sound realism by approximating the waveforms at the high frequencies and limiting them to 20 kHz with brick wall filters. Theoretically digital recording is fine, but the human ear deserves a higher sampling frequency. Perhaps only a select few can really hear the difference, but then how can we argue with them? In other fields, such as television, the trend is toward high-definition TV, in VCRs and camcorders there is a SVHS system, and yet tube-type audio amplifiers are still sold at premium prices because of many so-called golden ear audiophiles don't want to give up the tube sound. The same is with LP records. For the average listener, CDs are great as long as they don't hear pops and clicks and cannot break the stylus or the tonearm.

This chapter will discuss playback equipment. To understand the production of records/discs, refer to the *Handbook for Sound Engineers–The New Audio Cyclopedia* First or Second Edition.

27.2 Disc/Record Dimensions

The analog record has been standardized to 7 inch, 10 inch, and 12 inch discs and 33⅓ and 45 revolutions per minute (rpm).

Excerpts from the latest EIA standard for producing analog disc records are as follows.

27.2.1 Record Diameter

The diameter of records are:

12 inch LP disc, 33⅓ rpm	11.875 ± 0.031 inch (301.6 ± 0.8 mm)
10 inch disc, 33⅓ rpm	9.875 ± 0.031 inch (250.8 ± 0.8 mm)
7 inch disc, 45 rpm disc	6.875 ± 0.031 inch (174.6 ± 0.8 mm)

The recorded surface shall start with at least one turn of unmodulated groove.

27.2.2 Maximum Outer Diameter

The maximum outer diameter of a recorded surface shall be:

12 inch LP disc, 33⅓ rpm	11.500 inch (292.1 mm)
10 inch disc, 33⅓ rpm	9.500 inch (241.3 mm)
7 inch disc, 45 rpm disc	6.625 inch (168.3 mm)

27.2.3 Groove Dimensions

The groove dimensions shall be:

Minimum top width (monophonic only)	0.0022 inch (0.56 mm)
Maximum bottom radius	0.00025 inch (0.006 mm)
Included angle	90° ± 5°

On stereophonic records, the instantaneous groove width should be not less than 0.001 inch (0.025 mm). The average groove width should preferably be not less than 0.0014 inch (0.035 mm).

27.2.4 Stereophonic Groove

The *stereophonic groove* shall carry two channels of information. The two channels shall be recorded in such a manner that they can be reproduced by movement of a reproducing stylus tip in two directions at 90° to each other and at 45° to a radial line through the stylus tip and the center of the record. The reproducing stylus tip motion shall be tangential to, or lie in a plane through, the stylus tip and the record center, preferably inclined at an angle of 20 ± 5° clockwise to the normal to the record surface through the stylus tip, as viewed from the record center. In practice, angles of between 0° and 25° may be encountered.

27.2.5 Channel Orientation

The groove shall be recorded for reproduction with the right-hand loudspeaker(s), as viewed from the audience, actuated by movement of the groove wall, which is farther away from the center of the record.

27.2.6 Channel Phasing

The phasing of the two recorded signals shall be suitable for reproduction on equipment so connected that movement of the reproducing stylus tip parallel to the record surface (as with a monophonic record) produces in-phase signals across the output terminals of the phono cartridge.

27.2.7 Channel Levels

The levels of the two recorded signals should be such that peak excursions of the groove should not exceed 100 µm or 0.004 inch in lateral plane and 50 µm or 0.002 inch in vertical plane.

27.2.8 Speed of Rotation

Records shall be recorded for reproduction at one of the following speeds:

50 Hz Electric Supplies	60 Hz Electric Supplies
45.11 rpm ±0.5%	45.00 rpm ±0.5%
33⅓ rpm ±0.5%	33⅓ rpm ±0.5%

(Note: 16¾ rpm and 78 rpm speeds omitted.)

27.2.9 Lead-In Groove Pitch

The lead-in groove pitch shall be 16 ±2 lines/inch (l/in).

27.2.10 Lead-Out Groove

The pitch of the lead-out groove shall be 2–6 l/in. The top width of the lead-out groove shall increase to a minimum of 0.003 inch (0.076 mm) when the pitch exceeds ¼ inch (6.4 mm).

27.2.11 Finishing Groove

The diameter of the finishing groove shall be:

12 inch and 10 inch discs	4.187 ± 0.31 in (106.4 ± 0.8 mm)
7 inch discs	3.875 ± 0.078 in (98.4 ± 2 mm)

27.3 Signal Equalization in Disc Recording

To overcome the limitations found in the basic disc-cutting and reproducing process, special equalization of the signals before and after the recording was developed. When all signals that appear in the program bus are analyzed, we can see that the amplitude is the highest at low frequencies and the lowest at high frequencies. The relationship between the frequency of the signal and its amplitude where amplitude is inversely proportional to frequency is called a *constant velocity characteristic*, Fig. 27-1.

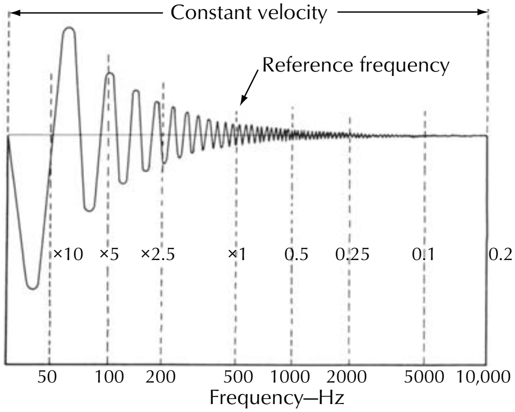


Figure 27-1. Constant velocity characteristics.

If the signals are recorded without equalization as they arrive, the low-frequency excursions would take all the space. The high frequencies would be of such a low amplitude that during the playback, high-frequency signals could be very close to the noise level of the system. The SNR then would be extremely small. This problem was recognized in the early days of disc recording, but the remedy used was only partial. At first only the low end of the audio spectrum was equalized. The cutting head sensitivity was decreased at low frequencies so that the amplitudes in midrange and at high frequencies could be recorded at higher levels. Then, the playback amplifiers were adjusted to boost the low frequencies to compensate for the losses introduced in recording. From this point on, the equalization used for cutting was called *preemphasis*, and equalization used in playback equipment, *postemphasis*.

The 1 kHz signal was chosen as the reference point because it was a convenient halfway point between the low and high frequencies. As time went by and further improvements were made, the equalization was extended to the higher frequencies as well. What emerged from the long and at times controversial subject of equalization are the RIAA and NAB equalization curves. The first curve was used by the Record Industry Association of America (RIAA) and the second, which is almost identical to the first curve, by the National Association of Broadcasters (NAB).

People still debate about two versions of the recording equalization. The DIN (Deutsche Industrie Norm) standard used in European countries calls for additional equalization at the extreme low end during playback to improve the SNR and stability of the system due to mechanical disturbances—i.e., turntable rumble—which can affect the overall performance of the system.

The NAB (RIAA) curve used presently in the playback equipment is shown in Fig. 27-2. The numerical values for the characteristic are shown in Table 27-1. For recording, the inverse curve is used. It means that if the playback signal is boosted +19.3 dB, the same signal should be recorded at the level of –19.3 dB so that the overall result will be 0 dB deviation from the ideal flat response curve.

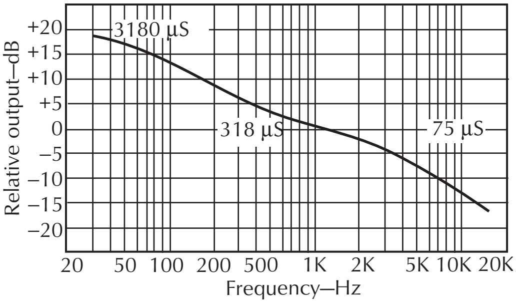


Figure 27-2. NAB (RIAA) standard reproducing characteristic.

Equalization is used to record the sound at the most advantageous levels for the best results as far as distortion and noise are concerned and to reproduce it so that the original balance between the frequencies can be restored. The RIAA curve is used for phonograph discs. Tape recorders record signals on tape, and tape recording has limitations that differ from the limitations found in mechanical recording and, therefore, require different preemphasis and postemphasis for best results.

Table 27-1. Preferred Frequencies and Calculated Recording Characteristics

Frequency (Hz)	Recording Characteristics (dB)	Frequency (Hz)	Recording Characteristics (dB)
20.0	–19.3	800.0	–0.8
25.0	–19.0	1000.0	0.0
31.5	–18.5	1250.0	+0.7
40.0	–17.8	1600.0	+1.6
50.0	–16.9	2000.0	+2.6
63.0	–15.8	2500.0	+3.7
80.0	–14.5	3150.0	+5.0
100.0	–13.1	4000.0	+6.6
125.0	–11.6	5000.0	+8.2
160.0	–9.8	6300.0	+10.0
200.0	–8.2	8000.0	+11.9
250.0	–6.7	10,000.0	+13.7
315.0	–5.2	12,500.0	+15.6
400.0	–3.8	16,000.0	+17.7
500.0	–2.6	20,000.0	+19.6
630.0	–1.6		

The RIAA curve covers the range from 20 Hz–20 kHz. The DIN curve, as shown in Fig. 27-3, extends the control over playback down to 2 Hz where the equalization returns back to 0 dB. As can be seen from the graphs, the curves have complex shapes; equalizer circuits use capacitors and resistors, and their values determine the amount of signal equalization that can be expressed as a function of a time constant in microseconds as derived from the equation

$$T = CR \tag{27-1}$$

where,

T is a time constant,

C is capacitance in farads,

R is the total effective resistance of the supply network in ohms.

This is part of the equation to determine the attenuation at various frequencies:

$$attenuation_{dB} = 10\log(1 + \omega^2 T^2) \tag{27-2}$$

where,

ω is $2\pi f$,

T is CR of Eq. 27-1.

The RIAA curve consists of three time constants; 75 µs to roll off the high frequencies, 318 µs to produce the slope below 1 kHz with a knee at 500 Hz, and a

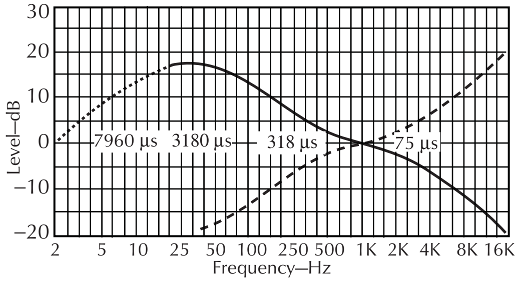


Figure 27-3. DIN recording and playback characteristics.

3180 µs time constant to flatten the low end of the curve. In today's modern amplifiers, the equalization is accomplished by placing the network with proper time constants into the negative feedback loop of the amplifier, thereby achieving lower distortion, better SNR, and improved signal-level-handling capability of the circuit.

Because the recording space on the record disc is limited, records are cut with constant amplitude characteristics of the signals in the upper half of the frequency range. When reproduced by the pickup, these signals are equalized to a constant velocity characteristic. In playing back these preemphasized disc recordings, different equalization has to be used for different types of cartridges. For instance, dynamic cartridges, which include moving-magnet, moving-iron, moving-coil, and variable-reluctance pickups, are constant velocity devices; therefore, they respond to the speed of the stylus movement. The faster the stylus is deflected, the higher the output voltage. Ceramic or crystal cartridges are pressure-sensitive devices, and they respond to the force applied to the stylus. They are called *constant amplitude devices*, and when records with constant velocity recording are played with ceramic cartridges, no additional equalization is required. The combined characteristics of both the recording and the cartridge complement each other, returning the signals to their original form. Only a minimal amount of signal grooming may be necessary to compensate for the effects of capacitive loading and nonlinearity of the cartridge.

27.4 Turntables

To play a record, the turntable or device to rotate the disc at the required speed is needed. This is the basic requirement for all turntables.The construction and execution of the requirement may differ greatly between the models and the designs of different manufacturers. The history of evolution of the record drive mechanisms takes us from the days of hand-cranked cylinder machines, through the age of spring-wound phonographs with mechanical governors for speed control, and into the age of electrically driven machines with electronic control. Today the accuracy of turntable speed is measured in small fractions of 1% in deviation from the desired speed.

27.4.1 Drive Systems

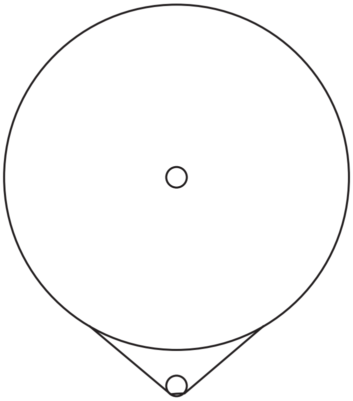
Turntables are driven by electric motors. The method by which the power from the motor is transferred to the turntable platter classifies the drive mechanism. The turntable platters can be belt driven, puck or idler driven, and driven direct.

The first category, the *belt-driven type*, encompasses all models that have motors mounted to the side of the platter with the belt stretched over the motor pulley and outer rim of the platter, Fig. 27-4A. Some platter designs have an additional internal rim to hide and to protect the belt.

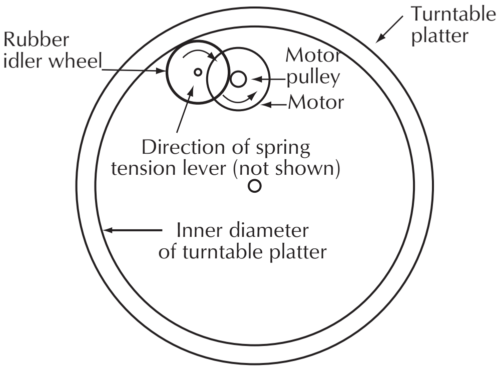
Many turntables have synchronous motors or motors with some type of speed control mechanism, such as a centrifugal switch that disconnects the power to the motor when the speed exceeds the preset value. The later types of motors are usually low-voltage, battery-driven motors used in portable equipment. Also, in portable turntables there is electrical feedback to control the speed of the low-voltage motor.

Another version of the same idea uses a low-voltage ac motor driven by a self-contained crystal-controlled oscillator allowing variation of the speed of the platter and achievement of great speed precision. The only source of speed variation can come from belt slippage or a defective belt. Belt-driven turntables are normally the quietest turntables. The speed selection of the belt-driven turntable can be accomplished either by changing the speed of the motor or by having the stepped pulley on the motor and by shifting the belt from one pulley onto another.

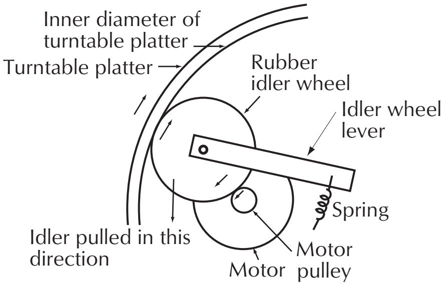
The second type of turntable is a *puck-driven* or *idler-driven turntable*, Fig. 27-4B. The coupling between the platter and the motor shaft is achieved through the intermediate idler wheel or puck, which has the outer edge covered with neoprene rubber or polyurethane for positive drive and to isolate the motor vibration from the platter. The idler wheel rotates on the shaft that is attached to a sliding bracket. When one side of the idler pulley (or puck) is in contact with the inner side of the rim of the platter and on the other side with the motor shaft, the idler wheel will transmit the motor rotation to the turntable platter. The mechanism is designed so when the motor is turned off the idler wheel



A. Single-motor belt-drive system.


Turntable platter
Rubber idler wheel
Motor pulley
Motor
Direction of spring tension lever (not shown)
Inner diameter of turntable platter

B. Puck- or idler-driven turntable.



C. Lever system jams the idler wheel between the motor pulley and the platter rim.

Figure 27-4. Various types of drive mechanisms.

retracts away from the motor shaft to protect the rubber ridge from forming a flat spot.

The advantage of the rim drive is that it provides positive torque to the platter, and if the motor is strong enough, it can bring the turntable to the desired speed almost instantly. The mechanism is simple, and it is the most reliable type of drive. Unfortunately it is also the noisiest because of the positive coupling between the motor and the platter idler or puck that transmits a certain amount of the motor vibrations to the platter and consequently to the record, as shown in Fig. 27-4C.

The third kind of turntable drive is the *direct drive* where the motor drives the shaft of the platter directly. There are also variations of the design. Some turntable designs are very sophisticated, using the platter itself as a rotor of the motor and drive is provided by the self-contained, quartz-controlled oscillator. The motion is extremely accurate and the speed of rotation may be displayed on the digital display, which is part of the control panel. There is also a weak point in this seemingly perfect drive. Because of the slow speed at which the turntable rotates, and because the motor has a finite number of poles, there is a slight cogging action in the platter motion, which may manifest itself with increased loads. This handicap is only related to turntable platters with fairly small mass and small moments of inertia. If the platter is heavy, it will overcome this problem.

The performance of the turntable depends very little on the type of drive used but more on the correct execution of the design by understanding the problems involved. The ideal turntable should have the following properties:

- It will start fast without hesitation.
- It will rotate with exact speed without variations.
- There will be no motor noises or vibrations heard while the system is in operation, they will not be transmitted to the platter.
- The turntable should be adequately shock mounted and isolated from the surface on which it sits to prevent the transmission of rumble and vibrations from the room. These loud sounds can actually shake the platter and the tonearm.
- The platter should be treated against ringing either by using a turntable mat with damping properties or by undercoating the platter.
- The turntable must be easy to maintain and to repair.

Not many turntables meet all these criteria; therefore, in order to know how to evaluate the unit, it is important to know how they work.

Speed of Rotation. Before evaluating the entire system, there are tests that can be performed on the turntable alone. The first one is speed of rotation. There are many ways of checking the speed of rotation, but the simplest one is by using the stroboscopic disc.

A *stroboscopic disc* is a circular disc containing a number of black-and-white bars, which are used for checking the speed of turntables and other rotating machines, Fig. 27-5. The disc is placed on the turntable, and the bars are observed under a fluorescent or neon light source fed from the normal ac lighting circuits. When the speed of the turntable is correct, the black bars appear to stand still. If the table is turning too fast, the bars speed up and drift in the direction of rotation. When running slow, the reverse takes place. Stroboscopic bars may be painted around the rim of a turntable and illuminated by a 115 Vac neon light mounted close to the table edge for constant observation. The equation for calculating the number of bars on a 60 Hz stroboscopic disc is

$$bars = \frac{2f60}{rpm} \tag{27-3}$$

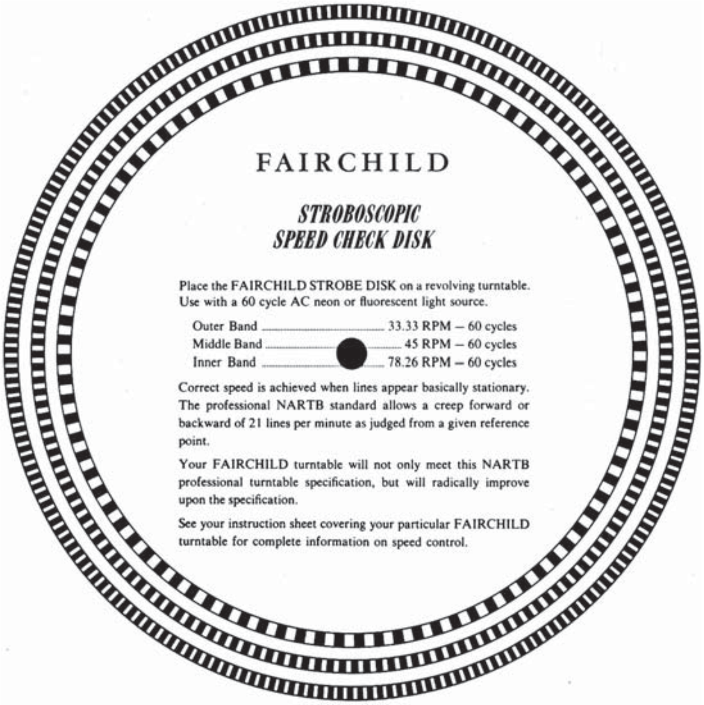
where,

f is the frequency of the strobe light used to observe the bars,

rpm is the speed of the turntable in revolutions per minute.

Starting Time. *Starting time* is the time it takes for the platter to reach its operating speed from a complete stop. This time period is important to know for professionals who have to begin playing the song or selection at the exact moment. To check the starting time requires either a stop watch or timing device and a strobe disc or the test record. As soon as the lines on the strobe disc appear stationary, the turntable has reached its operating speed. In playing the record test tone, the pitch changes as the correct speed is attained. Starting time may vary anywhere from a fraction of a second to two or more seconds, depending on the construction of the turntable. Turntables used by disc jockeys have to start as fast as possible without overshoot, which means that the speed should not, even for a moment, exceed the desired speed. If this overshoot occurs as the program material is already being transmitted, the variations of the speed will be most objectionable.

Acoustical Noise. The third test concerns the *acoustical noise* the motor and the turntable are producing. Normally, this test can be easily performed in a quiet listening room when everything is turned off and only the turntable is energized. If the turntable noise is clearly heard and it overshadows the normal room noise, turntable drive is below an acceptable performance level. A second part of the same test is conducted when the turntable is turned off and the system is



Place the FAIRCHILD STROBE DISK on a revolving turntable. Use with a 60 cycle AC neon or fluorescent light source.

Outer Band	33.33 RPM – 60 cycles
Middle Band	45 RPM – 60 cycles
Inner Band	78.26 RPM – 60 cycles

Correct speed is achieved when lines appear basically stationary. The professional NARTB standard allows a creep forward or backward of 21 lines per minute as judged from a given reference point.

Your FAIRCHILD turntable will not only meet this NARTB professional turntable specification, but will radically improve upon the specification.

See your instruction sheet covering your particular FAIRCHILD turntable for complete information on speed control.

Figure 27-5. A stroboscopic disc used for checking the rotational speed of a turntable. Courtesy Fairchild Recording Equipment Corp.

adjusted to a normal listening level. When the record with the quiet groove is placed on the turntable, a slight hiss can be heard when putting your ear to the loudspeakers. When the record with the quiet groove is placed on the turntable and the stylus is placed into the groove, listening to the increase in noise will show the extent to which the turntable transmits the building rumble. If the power to the turntable is turned on, the noise contributed by the motor drive can be measured. During this test, slightly tapping the base of the turntable can determine if the shock mounting is adequate and whether or not loud music will add coloration to the signal being reproduced. In summary, what is required from the good turntable is that it reproduces only what is recorded on the disc and is insensitive to all other sources of vibration.

27.4.2 Turntable Design in the 21st Century

One of the most important features of turntable design is the ability to keep noise and rumble created by motors and bearings from being picked up by the cartridge stylus. Many inexpensive turntables have a direct drive between the motor and the platter and inexpensive bearings, allowing motor noise and vibration to be transmitted to the platter and then to the cartridge stylus. Remember, it doesn't make any difference to the signal whether it comes from the stylus moving versus the disc or the disc moving versus the stylus.

VPI turntables, Fig. 27-6 use inverted bearings instead of conventional bearings. In this design the bearing assembly is in the platter rather than in a bearing well below the platter. The spindle and ball are attached to the chassis and the bearing well is inverted and placed in the platter itself. With this design the drive belt pulls through the center of the bearing assembly rather than many inches away from the center of the assembly, reducing teeter-totter effects to near zero for better stability.

All motor assemblies are completely separated from the turntable platter and tonearm, so there is no mechanical connection between the motor and the chassis except through the belt. This gives much lower noise levels due to isolation from the source of noise.

The VPI HR-X turntable uses a dual motor flywheel assembly to drive the platter. Two synchronous motors, driven by a perfect sine-wave ac power supply, drive a 14 lb flywheel spinning at 300 rpm, which in turn drives the platter. In this configuration the platter is driven by a non-electromotive source as opposed to other tables that are driven by the motor or combination of motors. Running the platter with no motor or multiple motors produces a velvety black background and perfect speed stability.

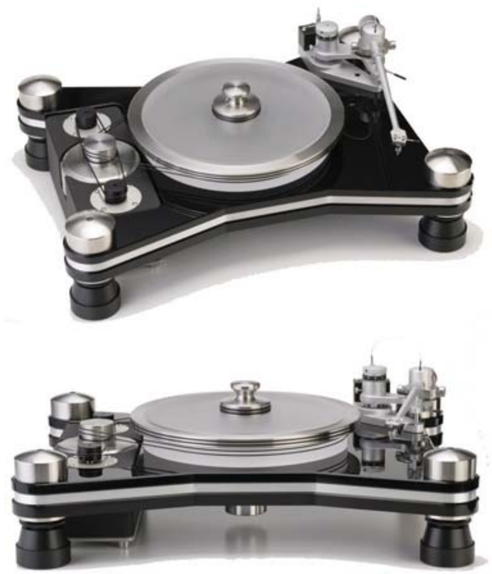


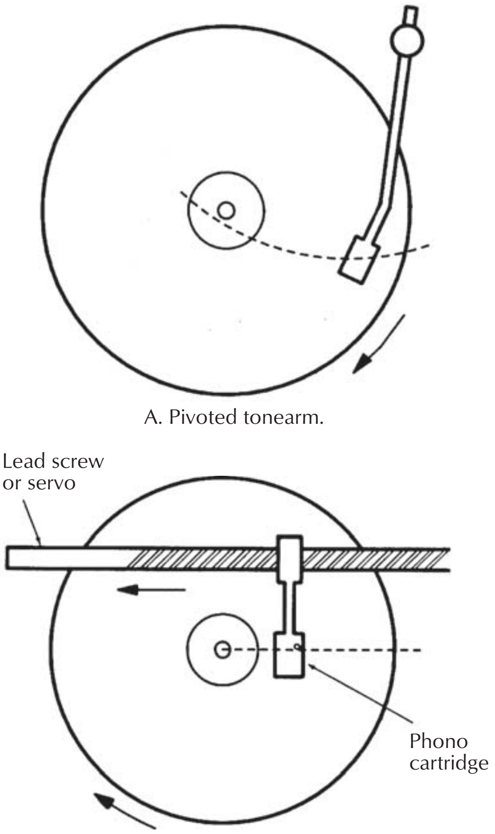
Figure 27-6. High-quality noiseless turntable. Courtesy VPI Industries, Inc.

27.5 Tonearms

Tonearms can be classified into two categories: *pivoted* and *tangential tracking*, Fig. 27-7A and 27-7B.

Contemporary tonearms are designed to cope with a variety of problems. However, rarely can one find a tonearm with nearly perfect geometry and correct design to establish correct performance. Most tonearms have built in antiskating devices, adjustable counterweights to accommodate a variety of cartridge weights and tracking forces, vertical height adjustment to set the tonearm parallel to the record, and a variety of features to facilitate installation and operation of the device. All tonearms are at best a compromise. Very few tonearms are dynamically balanced, and most rely on dynamic unbalance to produce vertical tracking force. The dynamically balanced tonearm is the tonearm that is capable of playing a record with the turntable tipped at almost any angle without changing the tracking force and tracking ability.

Tonearm Geometry. Tonearms are designed to retrace the modulation of the groove in the same way as it was recorded. Design of the tonearm takes into consideration the diameter of the records or the turntable, and the distance between the center of the platter and the



A. Pivoted tonearm.

B. Tangential tracking tonearm.

Figure 27-7. Tonearm classification.

pivot point of the tonearm. Older tonearms suffered from a tangent error because the cartridge was aligned properly at only one point on the record. Today's pivoted tonearms have a built-in offset angle at which the cartridge is positioned so it is always perpendicular within a couple of degrees to the radius of the disc. This reduces distortion in the lateral plane and improve tracking. There are many protractors available today using different approaches to help position the cartridge as accurately as possible in the tonearm to minimize tracking error.

When a disc is being cut, the cutting head is carried across the face of the recording disc following the radius. However, when in playback, the pickup is at the right angle to the radius of the disc only at two points, because the pickup arm is pivoted in such a manner that it swings across the face of the disc in an arc, as shown in Fig. 27-8.

Generally, the manufacturer of the arm supplies a template and mounting instructions for a particular arm. In the absence of such information, the pickup arm is mounted in such a manner that the tangent error is at a minimum. One method of mounting the arm is shown in Fig. 27-9. Regardless of where the pivoted arm is placed, a tangent error cannot be eliminated entirely.

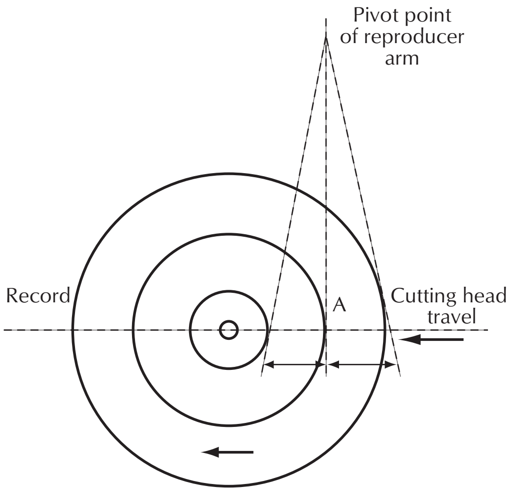


Figure 27-8. Tangent error in a reproducing arm. The error is zero at point A only.

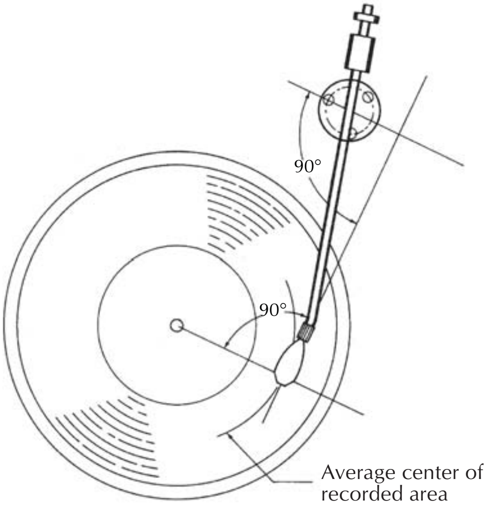


Figure 27-9. Typical mounting for an offset pickup arm.

The error can be made so small, however, that it can be neglected. In offsetting the tonearm by bending it into an S or J shape, Fig. 27-10, it is possible to position the cartridge so that at two points on the record the error shall be zero. The deviation from this ideal groove-cartridge interface will be only 2–3° in the horizontal plane.

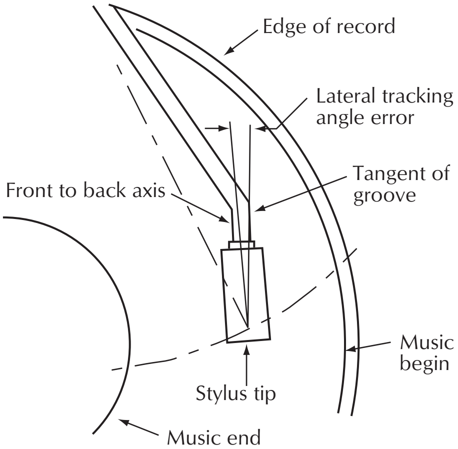


Figure 27-10. Geometry of a modern tonearm.

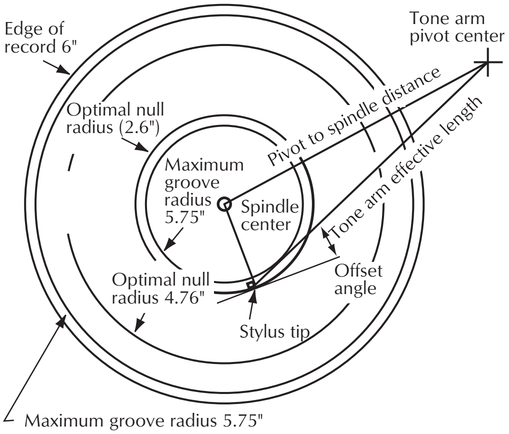

Edge of record 6"
Tone arm pivot center
Pivot to spindle distance
Tone arm effective length
Optimal null radius (2.6")
Maximum groove radius 5.75"
Spindle center
Optimal null radius 4.76"
Offset angle
Stylus tip
Maximum groove radius 5.75"

Figure 27-11. Relationship of the lateral components of a tonearm.

Offsetting the tonearm introduces the skating force that pulls the tonearm toward the center of the record. In tonearms without the offset angle the skating force is zero at one point and increases as the tonearm moves away from this position. The zero tangent error point in this tonearm coincides with the zero skating force position, point A in Fig. 27-8.

Theoretically, the pivoted tonearm without the offset angle and without any tangent error has to be infinitely long. The tonearm designed by the Rabinoff brothers revived the principle of tangential tracking used by Edison and found application in many turntable systems. In this system the tonearm motion has been achieved using servomechanisms and utilizing various types of arm position sensors. These tangential tracking turntables practically eliminated the tracking error and are quite popular with many hi-fi enthusiasts. There are also drawbacks to this design. Usually, such tonearms cannot be moved as fast as pivoted counterparts, and this may become a handicap in operations when speed of positioning the tonearm is of essence. The advantage of tangential tonearms is that they are shorter, lighter, and can be made more rigid to prevent many tonearm resonances found in some inferior pivoted tonearms. But the mechanical complexity of tangential tracking tonearms requires the use of modern technology including special integrated circuits and sensors.

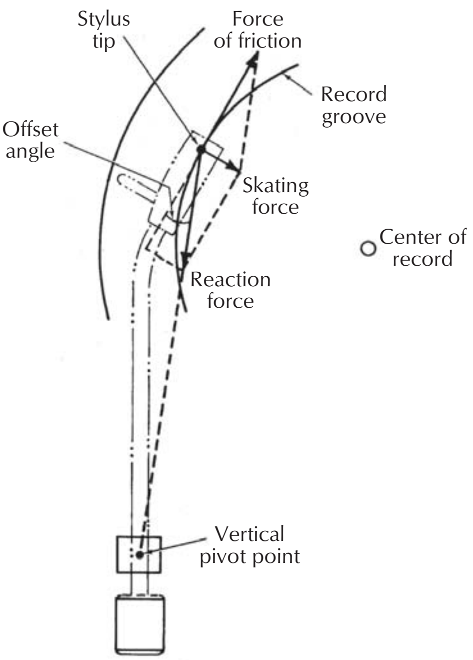
Effective Tonearm Length. Fig. 27-11 defines the turntable platter and spindle location in relationship to the effective tonearm length, which is the distance between the stylus tip and the tonearm pivot.

Modern tonearms have a built-in stop preventing them from moving farther than the locking groove so only three dimensions are of importance: effective tonearm length, vertical pivot-to-spindle distance, and the offset angle.

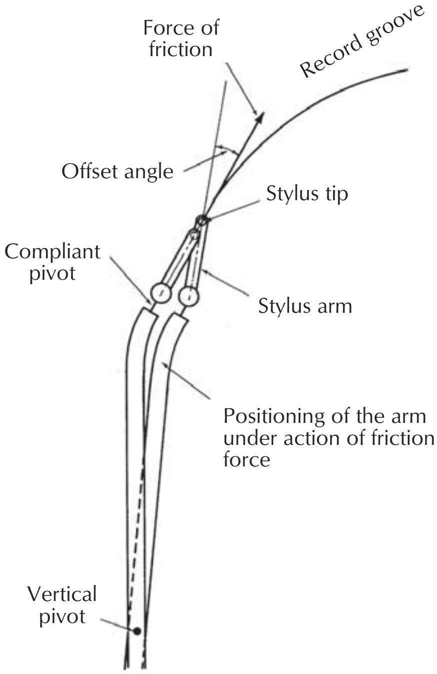
The accuracy of the cartridge tracking and mounting depends on the effective length of the tonearm. If the effective length of the tonearm is 7.87 inches and it is properly mounted (7.04 inches away from the turntable spindle), the cartridge will track to within +2¼° and –1½°, providing the cartridge is mounted at an offset angle of 27.8°. If the tonearm is longer, the lateral tracking error gets smaller so that the tonearm with the effective length of 10 inches will have a maximum tracking error of less than 1° at the smaller disc radius and a 1.7° error at the maximum radius.

Since the linear speed of the outer grooves is higher and the wavelengths are longer, tracking angle errors have lesser effect on the signal quality. Therefore, tracking errors should be minimized at the inner grooves for consistent quality of playback signal at all radii.

Skating Force. *Skating force* is a force that can upset the best aligned tonearm and cause considerable tracking error. The skating force is the result of tonearm geometry and the friction between the stylus and the record groove. Because of the offset angle and the overhang, one vector of this force pulls the stylus in a direction away from the pivot point of the tonearm and the second vector pulls the tonearm toward the center of the turntable, Fig. 27-12A. If this skating force is not compensated for, the stylus will be deflected toward the outside of the disc at the angle much greater than the error angle encountered in tracking the groove at different radii, Fig. 27-12B.


Stylus tip
Force of friction
Record groove
Offset angle
Skating force
Center of record
Reaction force
Vertical pivot point

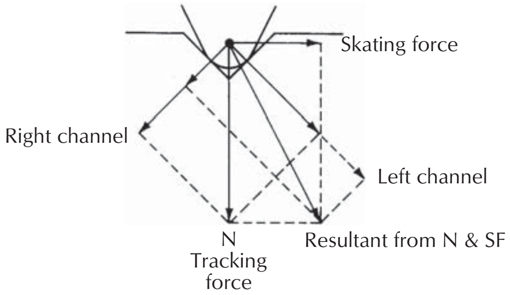
A. Position of the arm on the record and the forces acting on it.


Force of friction
Record groove
Offset angle
Stylus tip
Compliant pivot
Stylus arm
Positioning of the arm under action of friction force
Vertical pivot

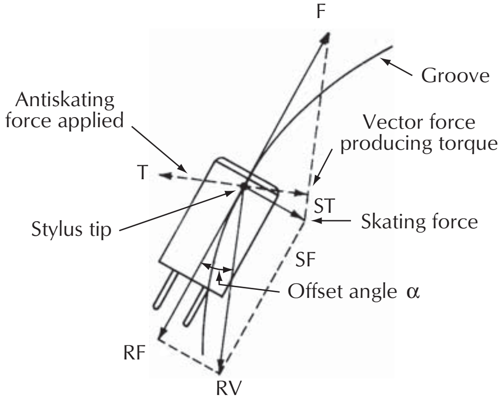
B. Effect of friction on tracking error.

Figure 27-12. Effects of tonearm geometry. Courtesy G. Alexandrovich.

The skating force compensation consists of applying a force to the tonearm that is equal to but opposite in direction to the skating force, Fig. 27-13. For all practical purposes, the skating force is constant for all radii of the music groove if the tracking error is small and the tonearm alignment is correct. There are slight variations of the skating force due to heavy modulation and groove wall plastic deformation caused by the sharpness of the new stylus, but the largest deviation in skating force is due to the variations in record material. From the study of various materials, it was established that the softest materials produce more friction and larger skating force. Lacquer masters produce up to 25% more friction (i.e., skating force) than vinyl records. Styrene records, today's 45 rpm discs, have approximately 30% less friction than vinyl, requiring less antiskating compensation than vinyl LPs.


Skating force
Right channel
Left channel
N
Tracking force
Resultant from N & SF

A. Rear view.


F
Groove
Antiskating force applied
Vector force producing torque
T
ST
Stylus tip
Skating force
SF
Offset angle α
RF
RV

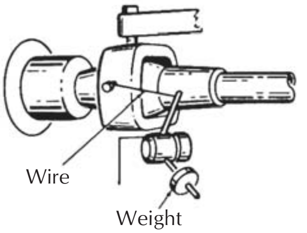
B. Top view.

Figure 27-13. Skating and antiforces in a record groove. Courtesy G. Alexandrovich.

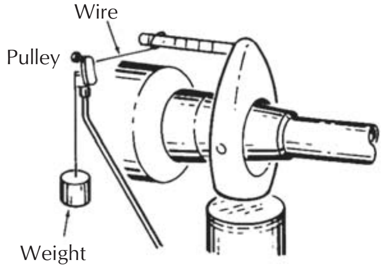
There are many different ways to generate the antiskating force. It is incorrect to assume that increasing the drag on the horizontal motion of the tonearm will compensate for skating. Skating force is independent of groove spiraling speed; drag is not. Also, because of the variable pitch common to all present-day recordings, the speed with which the tonearm moves across the record varies and at times may even be zero. Because of this variation, the mechanism that generates the antiskating force should be able to generate a uniform force at all times, regardless of the motion of the tonearm. Antiskating force can be generated by using springs, magnets, weights with pulleys, electrical devices, and mechanical linkages and weights, Fig. 27-14. Any method to apply the clockwise bias in a horizontal plane to the tonearm to counteract the skating force produces positive results; however, compensation may not be accurate for all types of systems.

The effectiveness of the antiskating force mechanism depends to a high degree on the dynamic behavior of the tonearm. If the tonearm is not dynamically balanced (and most of them are not), any tilt of the turntable may result in a change of skating force, endangering the tracking ability of the pickup. As was previously mentioned, the dynamic balancing of the tonearm implies that the pivot point of the tonearm is also the center of mass. In most modern tonearms this center of weight is shifted toward the cartridge end in order to produce tracking force, Fig. 27-15. In a dynamically balanced tonearm, tracking force is produced by using either a spring or a permanent or electromagnet (solenoid). A properly dynamically balanced tonearm could play a record with the turntable being in any position and is completely insensitive to jarring of the turntable or floor vibrations.

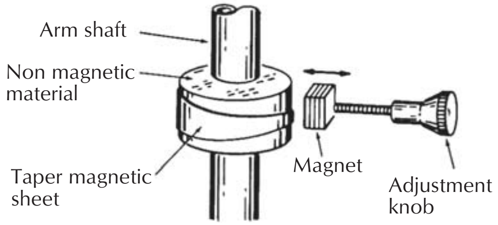
Vertical Tracking Angle. An important adjustment of the tonearm is in positioning the cartridge over the surface of the disc. Cartridges are mounted in tonearms so that the mounting surface of the cartridge is parallel to the record surface, Fig. 27-16A. Sometimes tilting the cartridge fore or aft results in lower tracking distortion. Some cartridges are designed to produce the lowest distortion when playing vertical modulation that was recorded at the vertical cutting angle of 25°, Fig. 27-16B. At the same time most of today's records are cut with the vertical angle of 10°–15°. So in order to reduce the distortion during playback, matching the two angles by moving or tilting the cartridge backward a few degrees may help reduce tracing distortion.



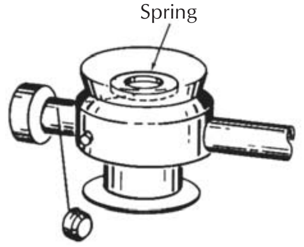
A. Mechanical linkages and weights.



B. Weights and pulleys.



C. Magnets.



D. Springs.

Figure 27-14. Different methods of generating antiskating force.

Tonearm Resonance Damping. A Shure Brothers, Inc. study revealed that the warp frequencies of LP records lie in the region from one revolution (0.5 Hz) peaking at 3 Hz and tapering down at 7–8 Hz. Because the audible range of frequencies starts at around 20 Hz, tonearm resonance placed between the warp frequency region and the audible region will allow minimum distortion of the signal due to tonearm bounce. As a result of this research, improvements were made in the tonearms by applying vertical damping to the tonearm. The vertical tonearm motion control was attacked by Discwasher, Inc., by designing a special damping mechanism named Disctracker, which attached to the cartridge. Shure Brothers introduced their stabilizer brush that attached to the cartridge similar to the brushes invented and used

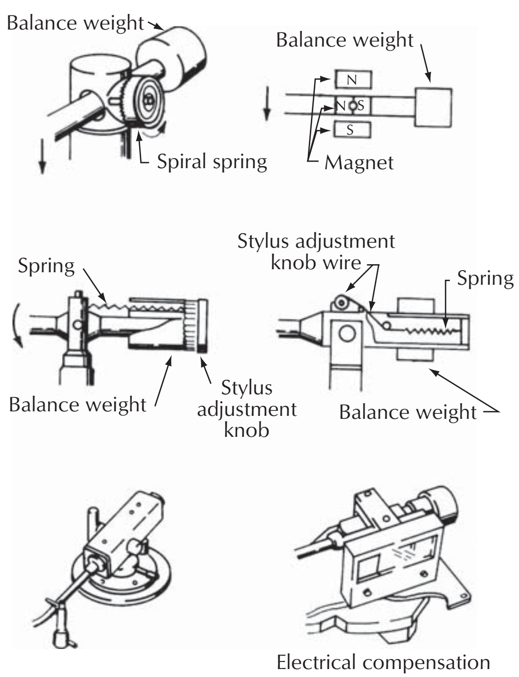
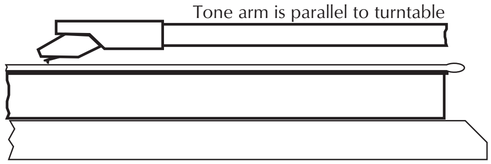
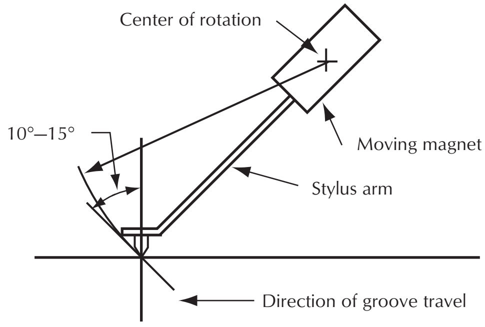

Balance weight
Balance weight
N
N S
S
Spiral spring
Magnet
Stylus adjustment knob wire
Spring
Spring
Balance weight
Stylus adjustment knob
Balance weight
Electrical compensation

Figure 27-15. Examples of how dynamic balance of the tonearm can be achieved.


Tone arm is parallel to turntable

A. Tonearm is parallel to turntable.


Center of rotation
10°–15°
Moving magnet
Stylus arm
Direction of groove travel

B. Cartridge designed to produce lowest distortion when playing vertical modulation.

Figure 27-16. Schematic representation of the moving system of a pickup, illustrating the vertical tracking angle.

by Pickering and Stanton since 1971, except that the Shure stabilizer brush had its pivots filled with damping fluid. These devices helped to various degrees to stabilize the tonearm as the brush cleaned the record groove.

The other approach was to adjust the effective mass of the tonearm by pivoting only the front part of the tonearm and selecting a cartridge with compliance that would match the mass of this portion of the arm, Fig. 27-17. Dynavector tonearm is an example of such design. Another variation is the design by Sony that employs electronic control of the tonearm motion. Instead of relying on weights, springs, or magnets, the Sony tonearm uses linear dc electromotors driven, operated, and controlled by electrical signals. Unfortunately, not all functions of the tonearm are controlled automatically and are subject to misadjustment.

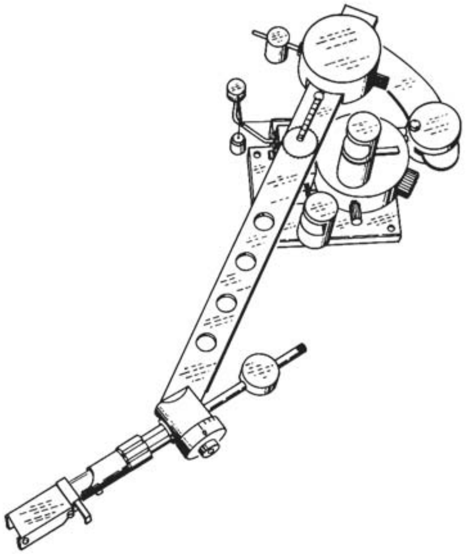


Figure 27-17. Dynavector tonearm with pivoted front portion for lower dynamic tonearm mass. Courtesy Onlife Research, Inc.

27.6 Phono Pickup/Transducers and Styli

In order to reproduce signals recorded on the phonograph record, a *transducer* (phonograph pickup, phono cartridge, or needle) converts the groove modulation into the electrical signals. Unlike microphones, loudspeakers, and other types of devices or transducers that convert one form of energy into another, the phonograph pickup has to perform more than one function. The phonograph pickup or cartridge, so called since the invention of the removable stylus or needle assembly, has to convert modulation of the record groove into the electrical signals, and at the same time support the tonearm at the proper height above the record surface, all the while moving the tonearm across the surface of the record.

The phonograph cartridge is an electromechanical device designed to track or follow the excursions of the record groove and to convert this motion, with the help of a tracking mechanism-stylus assembly, into electrical signals.

Cartridges are classified by the principle by which they convert mechanical motion into the electric current or signal, *electrodynamic* and *piezoelectric*. There are also pickups designed to operate using strain gauges, variable capacity, and light as sensors.

Electrodynamic-Type Cartridges. Electrodynamic-type cartridges are subdivided into three categories: *moving magnet*, *moving coil*, and *induced magnet* or *moving-iron* type. The electrodynamic principle consists of using a magnetic field that, when it intersects the coil windings, generates electric current. The construction of the cartridge classifies the type. If the magnet is attached to the stylus tube or cantilever and the coils are stationary, it is called a moving-magnet cartridge. If the magnet is made stationary and the coils move in the magnetic field, it is a moving-coil cartridge; and if the magnet and the coils are made stationary and there is a slug of soft magnetic iron moving in place of a magnet while being magnetized by the stationary magnet, it is called a moving-iron or induced-magnet cartridge.

Variable-Reluctance Cartridges. Since the introduction of the original variable-reluctance pickup, Fig. 27-18, many different versions of its design have appeared. The magnetic structure consists of two pole pieces A, with a small permanent magnet B between them. The coil C is mounted with a soft rubber insert D.

The stylus, which is also the armature, is held in the exact center of the magnetic structure by the rubber insert. When the stylus is actuated, its movement causes a voltage to be generated in the coil. Because of its construction, the frequency response extends beyond the normal audio-frequency band. Output voltage is on the order of 100 mV at 1 kHz, with an output impedance of 500 Ω. The recommended stylus pressure is 15–20 g. The stylus weighs 31 mg and is removable. Although the recommended pressure is 15–20 g, the pressure could be as low as 7 g. The frequency response is ±2 dB, 20 Hz–20 kHz.

Moving-Coil Cartridges. Modern moving-coil cartridges are represented by a variety of designs. All of them have coils that move, but not all of them are entitled to be called moving-coil type. Designs that depend for their functioning on the motion of the soft iron core rather than on the motion of the coil itself should not be classified as a pure moving-coil device. There the motion of the coil is coincidental. Fig. 27-19 shows the cross sections of moving-coil stylus assemblies as they move during the playing of the record. The magnetic flux is directed by the iron core or armature of the coil. If the coil is made stationary and the core is vibrated, the signal will still be generated. This fact prevents it from being classified as a pure moving-coil device.

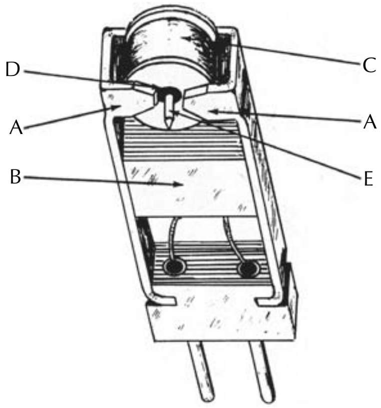
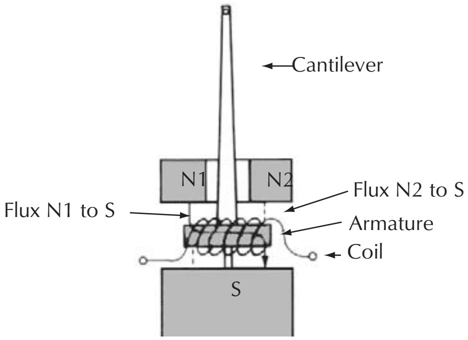


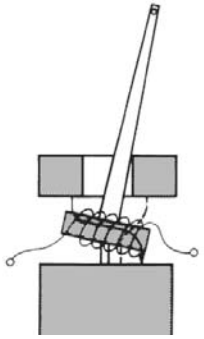
Figure 27-18. Variable-reluctance magnetic pickup.

The advantage of this design is extremely low output impedance, making the cartridge insensitive to capacitive loading and allowing the use of very long cables without altering the frequency response of the device. On the negative side, the output of the cartridge is very low, measuring in the tenths of a millivolt requiring an extra 20–30 dB amplification to bring the electrical signal to the required level as referenced to an established sensitivity of 1 mV/1 cm/s of recorded velocity. A step-up transformer or an extra stage of amplification usually introduces additional noise and the effect on capacitive loading. Other drawbacks of such design are the weight of the cartridge and the need to use a heavier tracking force.

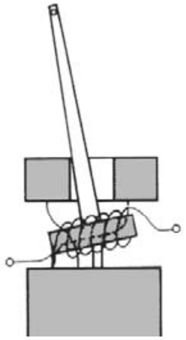
One of the debatable points about MC cartridges is the sound they produce. MC cartridges have a very fast response to transients because of the very low inductance and the impedance of the coils and the very rigid cantilever, which has to be strong in order to move a relatively heavy coil assembly. Another factor in this type of design is the construction of the coil assembly, which may have a number of turns in the coil unsupported and free to vibrate, producing random signals at higher frequencies. Also lead-in and lead-out wires may not be secured properly and can vibrate in the magnetic field producing random coloration of the signal. Lead



A. Cantilever in the neutral position.



B. Cantilever deflected to the right.



C. Cantilever deflected to the left.

Figure 27-19. Bottom view of the moving-coil cartridge generator assembly.

dressing, coil impregnation, and gluing techniques control the purity of the sound produced by this design.

Step-up transformers require winding ratios of 1:10 or more. The transformer's high-impedance secondary winding is reflected back into the primary, and any loading of the secondary in excess of the specified value affects the signal output level and electrical damping of the coils. Theoretically, shorted coils produce maximum damping, while an unterminated winding of the transformer's secondary will emphasize electrical resonances and unchecked mechanical motion. It is important to locate the step-up transformer near the preamplifier input to minimize the capacitive load of the shielded wires between the transformer and the input stage of the amplifier. Because the levels handled by this input transformer are extremely low, good transformer shielding is necessary.

In lieu of the step-up transformer, a prepreamplifier may be used. Additional preamplification, obtained from active gain circuits, requires super low-noise circuits in order to preserve an acceptable SNR. There have been many such pre-preamplifiers designed using the most exotic devices and circuits, operating with batteries or special ac power supplies with maximum filtering and voltage regulation and using magnetic shielding.

Moving-Magnet Cartridges. The most popular high-performance stereo cartridges are the moving- magnet type. Moving-magnet cartridges offer one of the most sensible ways to design the stereo cartridge with a replaceable stylus. This cartridge has low dynamic tip mass, high compliance, and fairly high output. By using the most powerful rare earth magnets and using the most modern manufacturing methods, the frequency response is extended from almost direct current to well past the threshold of hearing, Fig. 27-20.

Induced-Magnet Cartridges. An example of an induced-magnet or variable-reluctance pickup is manufactured by Bang and Olufsen of Denmark. It consists of a small armature in the form of a cross, made of Mumetal, which swings between four pole pins, Fig. 27-21A. A stylus bar constructed of aluminum tubing 0.002 inch (0.05 mm) thick is attached to the Mumetal armature cross at one end. The stylus is secured to the other end of the tube. Four pole pins with four coils are placed at each end of the cross. With a 45° motion to the right, a reverse voltage induction takes place. Such action permits the coils to be connected push-pull, thus reducing harmonic distortion induced by the nonlinearity of the magnetic field. In addition, the coils provide an effective hum-bucking circuit.

Crosstalk between the left and right channels is minimized, since such components are bucked out. Modulating one channel 45°, the cross arms on the orthogonal channel rotate without changing the spacing; therefore, there is no induced voltage in this channel, assuming the positioning of the unit, with respect to the groove, is correct.

A cross-sectional view of the magnetic circuit is shown in Fig. 27-21B and is similar to the magnetic structure of a loudspeaker employing a center magnet. Thus, a closed magnetic circuit, which prevents leakage

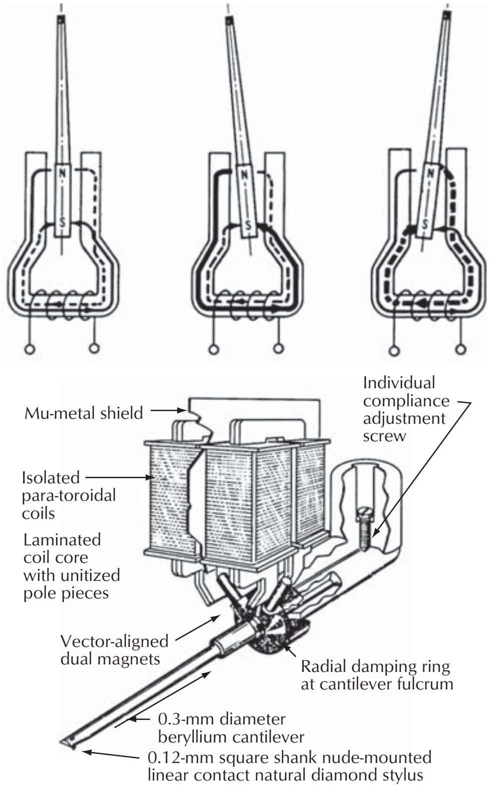
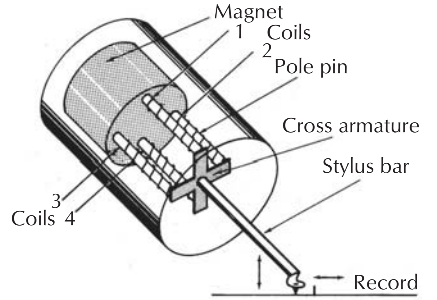
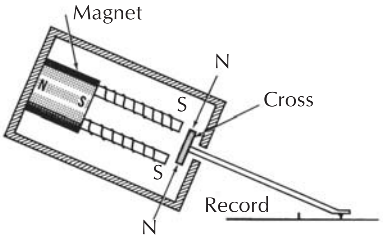

N
S
N
S
N
S
Mu-metal shield
Individual compliance adjustment screw
Isolated para-toroidal coils
Laminated coil core with unitized pole pieces
Vector-aligned dual magnets
Radial damping ring at cantilever fulcrum
0.3-mm diameter beryllium cantilever
0.12-mm square shank nude-mounted linear contact natural diamond stylus

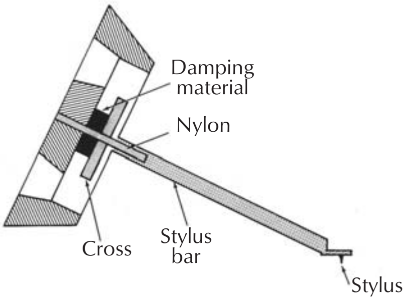
Figure 27-20. Basic principle of the moving-magnet pickup.


Magnet
1 Coils
2
Pole pin
Cross armature
Stylus bar
3
Coils 4
Record

A. Simplified drawing of coils and cross armature.


Magnet
N
S
N
S
Cross
S
Record
N

B. Magnetic circuit.


Damping material
Nylon
Cross
Stylus bar
Stylus

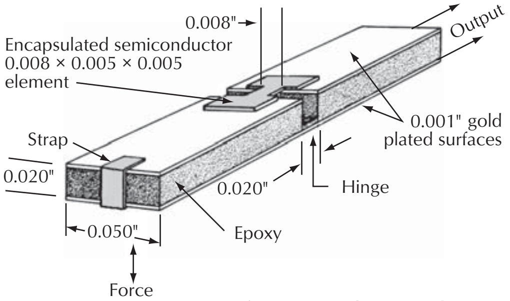
C. Stylus mounting with damping control.

Figure 27-21. Induced-magnet cartridge construction.

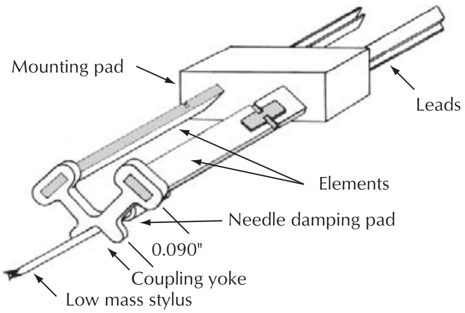
of the magnetic field, is provided and being nonmagnetic, it cannot be attracted to the steel turntable plate. It also provides an effective shield for the coils. The stylus bar pivots on a nylon thread, bonded to a plastic support. The armature cross bears on a resilient disc, Fig. 27-21C, which controls compliance and supplies damping for the moving system. The rotational point of the system is at the junction of the armature cross and the nylon thread support. The output voltage is 7 mV for each channel for a 5 cm/s cut. The stylus has an angle of 15° at 2 g of tracking force and may be operated at a pressure of 1–3 g. Compliance is 15×10^{-6} cm/dyn for both directions of motion. Frequency response is 20 Hz–20 kHz ±2.5 dB.

Semiconductor Pickup Cartridge. A *semiconductor pickup cartridge* operates on the principle of the strain gauge. The pickup mechanism employs two small, highly doped silicon semiconductor elements 0.008 inch × 0.005 inch whose resistance varies as a function of the stylus deflection, Fig. 27-22. The elements are mounted on laminated beams of lightweight epoxy with gold-plated surfaces. A notch in the beam under the assembly acts as a hinge for stress concentration. In this structure, two beams are used, each driven by an elastic yoke, coupled to the stylus. Aside from the compliance of the yoke and mounting pads, a mechanical advantage of over 40:1 can be attained in the beam and stylus lever. This mechanical transformer provides high compliance and reduces the mass of the elements reflected to the stylus. The stylus is elliptical in shape and set at an angle of 15°.

Since the semiconductor elements are sensitive modulating devices and not generators as in the conventional pickup, very little energy is required for their operation. The compliance at 1 kHz is approximately 25×10^{-6} cm/dyn and the frequency response is from 20 Hz–50 kHz. A power supply, two single-stage preamplifiers, and one inverter stage are required. As the elements are deflected by the stylus action, the resistance of the semiconductors, about 800 Ω, changes slightly, causing a varying dc voltage across the output. This dc signal is ac coupled to the preamplifiers in the power supply, providing an output voltage of 0.4 V for each side. The cartridge employs mechanical equalization that, in combination with the RC equalizer at the output of each preamplifier, results in an RIAA reproducing characteristic.



A. Beam construction of a semiconductor pickup.



B. Cartridge construction of a semiconductor stereophonic pickup.

Figure 27-22. Stereophonic semiconductor pickup.

Piezoelectricity. Piezoelectricity is pressure electricity. The voltage generated by the crystals in piezoelectric cartridges is proportional to the amplitude of the stylus displacement. The output voltage of the average piezoelectric pickup is considerably higher than for other type pickups. Piezoelectric pickups are treated electrically as a capacitive-reactance device since the impedance rises with a decrease of frequency. Simple RC networks are used with this type of pickup to obtain a frequency response corresponding to the standard RIAA reproducing characteristic. Records recorded using a constant-amplitude characteristic may be reproduced without equalization.

In the ceramic stereophonic pickup, Fig. 27-23, the moving system consists of two piezoelectric crystal slabs of lead-zirconium titanate or similar material. This particular material offers good mechanical and electrical properties with high sensitivity and high capacitance. The ends of the slabs are mounted rigidly in a mounting block, and the front end is connected by a yoke made of injected molded plastic. This coupling is critical because the electrical performance and the mechanical impedance seen at the stylus point by the record groove depends on it. The coupling system is defined as that portion of the mechanism that lies between the stylus tip and the ceramic slabs.

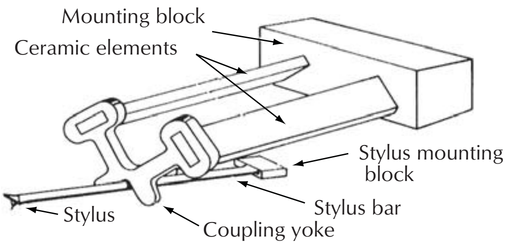


Figure 27-23. Simplified drawing showing the construction of a ceramic stereophonic pickup.

The stylus bar is made from heat-treated, thin-walled aluminum alloy tubing, with one end flattened to hold the stylus at the desired angle. The other end of the stylus bar is held in place by the stylus mounting block. The coupling yoke is connected at a point about midway on the stylus bar. This point is chosen because it affords the most desirable electrical performance and substantially reduces the mechanical impedance of the yoke and ceramic elements as seen by the stylus tip.

Better designs have four output terminals, two for each channel, to ensure the complete isolation of one side from the other. Damping in the form of a viscous material is used to control the frequency characteristics. These pickups are of the constant-amplitude type with the output voltage 10 mV for a peak velocity of 5 cm/s. Ceramic pickups are not affected by magnetic or electrostatic fields.

RC equalizer networks for both crystal and ceramic pickups are shown in Fig. 27-24. The networks are connected between the output of the piezoelectric pickup and the input of the preamplifier. The characteristics of these networks are such that they correspond to the standard RIAA reproducing curve. Using a pickup with a compliance of 15×10^{-6} cm/dyn or greater, the response can be within ±2 dB.

The internal impedance of the average crystal pickup is approximately 100 kΩ, with a capacitance of 0.001 to 0.0015 μF.

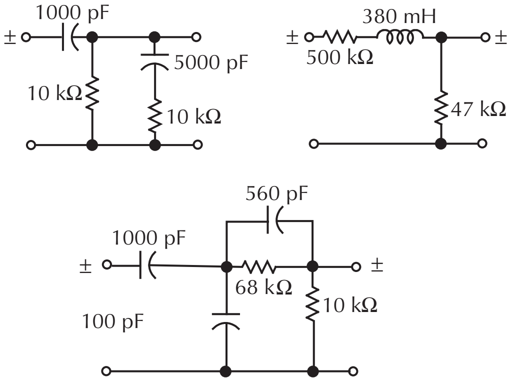


Figure 27-24. Networks for equalizing ceramic cartridges.

27.6.1 Cartridge Styli

Stereo Disc Groove. The playback stylus is the first link between the information stored in the record groove and the playback system. The quality of the reproduced sound is influenced by the precision with which the stylus follows the groove modulation.

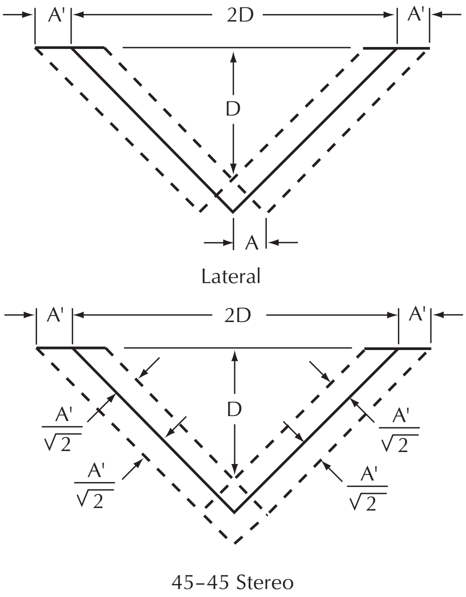
In stereophonic recordings with 45°/45° modulation, the two channels are isolated from each other because modulation of each channel is at 90° to the other, Fig. 27-25.

To minimize the effects of vertical excursions at low frequencies, the phase of both channels is adjusted so low-frequency signals are in phase in order to produce lateral modulation. The phase relationship of the two channels determines the location of the sound image between the two loudspeakers, and in some cases the phase is a deciding factor as to whether there is going to be a signal reproduced at all.

Stylus Tip. The function of the playback stylus of the cartridge is to follow all deflections of the groove. Since the stylus is attached to the end of the cantilever, any motion of the stylus tip is transmitted to the other end of the tube or shank, where the electrical signals are generated by a moving magnet, a moving coil, or a crystal. The stylus has rounded off edges that are polished for smooth tracking. Ideally, the playback stylus should be centered in the groove, and its centerline should match that of the cutting stylus. There are always minute imperfections in the alignment of the stylus and of the groove. Therefore the shape of the playbacd stylus is made to compensate ans allow some misalignment of the stylus in the record groove.

The stylus touches the groove walls at two points. The contact area is curved and is a part of the tip radius so that if the stylus is slightly tilted due to misalignment of the cartridge or the tonearm, tracking will not be affected.

Spherical Stylus. There are several types of styli today. The simplest and the oldest one is the *spherical tip*. The spherical stylus is a tiny diamond or sapphire cylinder with one end ground to a cone shape with its tip polished to an accurate sphere. The included angle of the cone is about 55°, and the tip radius is about 0.0007 inch or 0.7 mil. Because grooves can be as narrow as 0.001 inch, the stylus tip has to be equal to or smaller than the groove in order to track it. The standard tip radius dimensions for today's spherical styli range from 0.0005–0.0007 inch (12.7–17.7 µm).



Lateral		45-45 Stereo
2D + 2A'	Maximum groove excursion	2D + 2A'
A	Relative output per channel	$\frac{A'}{\sqrt{2}}$
0	Relative dB per channel	−3.0A'

Figure 27-25. Comparison of 45°/45° stereophonic groove with standard lateral groove.

Elliptical Stylus. The second type is the *elliptical stylus*. From the front it looks like a spherical stylus; however, there are two flats polished in the front and the back of the stylus. The side radius of the elliptical tip is much slimmer than that of the spherical stylus. The intersections of the two flats are polished to form small radii called the *tracing radii*, which measure about 0.0002 inch (5 µm). These small side radii are actually in contact with the modulation of the groove and, because they are small, they follow the high-frequency excursions of the groove more easily.

Stylus Characteristics. All playback styli are designed to contact only the walls of the groove; therefore, the stylus tip has to ride without touching the bottom of the groove. Since the diamond gets slimmer as it wears down, the tip gets closer and closer to the bottom of the groove. When it starts touching it, noise increases because debris has accumulated on the bottom of the groove and is scooped up by the stylus. This is a clue to change the stylus in order to reduce the noise and to preserve the record from being destroyed by the sharp edges of the worn diamond.

Currently, almost all styli manufactured are made out of diamond. The quality and the price of the stylus depends on whether it is made out of a solid piece of diamond or a small chip bonded onto another material that acts as an extension or pedestal for the diamond tip. The technology of manufacturing diamonds has advanced significantly so that chip bonding and encasing can be favorably compared to solid or nude diamond tips. In view of the fact that the area of contact is only 0.2 millionths of a square inch (0.2×10^{-6} inch) and as long as this area is made out of a diamond, the overall performance of the stylus will not be affected. All this is true providing that the mass of the bonded stylus assembly is not higher than that of a conventional diamond and not larger than the nude stone.

The vertical tracking force applied to the stylus is divided between the two walls. Each wall is experiencing force that is equal to the total vertical force times the cosine of 45° or 0.707, Fig. 27-26. For instance, if the vertical tracking force (VTF) is 1 g, each groove wall will experience a force of 0.7 g.

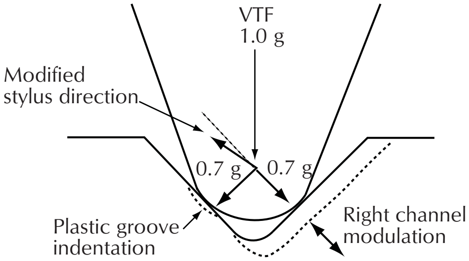


Figure 27-26. Stylus motion and forces acting upon it in a stereo groove.

A very small area of contact exists between the stylus tip and the groove so the pressure against the groove wall can rise up to many thousands of pounds per square inch. For instance, if each wall receives 0.7 g of force applied through the contact area equal to two ten millionths of an inch (0.2×10^{-6}), the pressure is 7726 lb/inch. With such high pressures and force of friction between the stylus and the vinyl, the outer skin layer of the record material melts as the tip slides over the plastic and then refreezes almost as fast as it melted. Since the melting temperature of the vinyl is about 480°F, the same temperature exists in the contact area.

Stylus Cantilever. The stylus is attached to some type of coupler or cantilever that connects it to the generating element of the cartridge, which could be a magnet, a piece of iron, a coil, or a ceramic element. Because of a very wide range of frequencies this stylus assembly has to transmit, the construction material and shape of the cantilever are very important. Theoretically, it has to be very light and rigid. Over the century of existence of mechanical sound recording, styli were made out of cactus needles, whale bones, and all kinds of metal, gems and stones, plastic, and wood. The final choice is centered around an aluminum alloy thin-wall tube. It is fairly strong, light, noncorrosive, nonmagnetic, electrically conductive, and easy to manufacture.

The average diameter of the aluminum cantilever tube is 0.03 inch (0.76 mm), and the length may vary from ¼–½ inch (6–12 mm). A few exotic cartridges have cantilevers made out of solid ruby or even diamond and some from boron or beryllium copper alloy. Although ruby and diamond are extremely rigid materials, because of manufacturing difficulties and high weight/length ratio, they are made very short. This, in turn, brings the pivot point much closer to the stylus tip that moves in a much smaller arc when reproducing groove modulation. Since the grooves are modulated by the cutting stylus that has its pivot quite a distance away and is moving in an arc of much larger radius, the larger the difference between the motions of the cutting and of the playback styli, the larger the distortion.

On the other hand very long playback cantilevers are unable to produce sufficient motion of the generating element that results in a very low electrical output.

Compliance. The amount of force required to move the playback stylus depends on several factors; the first is the compliance of the stylus, and the second is mass.

Compliance of the cantilever or the stylus is the ability of the stylus assembly to react to the groove modulation. It is measured in cm/dyn or µm/mN (metric) and gives the amount of stylus tip deflection for the given force. Compliance is measured statically and dynamically.

Static compliance is the amount of deflection of the cantilever when a constant force is applied to the stylus tip. Dynamic compliance is a measure of tip deflection as it is reproducing the frequency of known amplitude at which the measurement is being made.

Vertical Resonance. The second variable in the equation is the tonearm/cartridge vertical resonance. Tonearms and cartridges resonate between 5 Hz and 15 Hz; the most desirable range is between 8 Hz and 12 Hz. Resonance below 8 Hz will produce instability of the tonearm and will result in poor tracking of moderately warped records.

Stereo cartridges have fairly uniform compliance in all planes of stylus motion. Cartridges with higher compliance work best with light tonearms, and heavy tonearms should be set up with cartridges having low compliance. If the stylus compliance is low, the tracking force applied to the stylus should be higher than for a high-compliance stylus.

27.6.2 Cartridge Voltage Output

The output voltage of the cartridge depends on its design and the type of generator system used. Ceramic or crystal cartridges produce the highest voltage. Next are the moving-magnet cartridges and then the induced-magnet pickups; the last group is the moving-coil cartridges. The moving coil cartridge produces higher power output than other types so they can work with step-up transformers to increase the output voltage 10–20 times or 20–26 dB. On the other hand, some high output voltage ceramic cartridges are connected to the loss pads and response-shaping networks to reduce the voltage down to the average output level of the moving-magnet cartridges. Today most of the preamplifiers are designed to accept moving-magnet cartridges.

27.6.3 Electrical Loading

With various output levels and different source impedances, cartridges respond differently to electrical loads. For instance, crystal or ceramic cartridges are the most susceptible to capacitive loading. The entire frequency response is dependent on the loading of the cartridge. In the moving-magnet cartridge, only the highest portion of the frequency range is affected by the capacitive loading. The moving-coil cartridge is almost completely immune to the loading effects. Once it is connected to the step-up transformers, the secondary of the transformer becomes very sensitive to loading, and excess capacity can play havoc with transformer resonance and the impedance of the secondary transformer winding. Therefore, cartridge manufacturers specify the recommended resistive and capacitive loads.

The most common resistive load is 47 kΩ (50 kΩ for Europe), paralleled by 200–400 pF of capacitance for the moving-magnet cartridges, depending on the manufacturer and on the cartridge model. The capacitive loading for the cartridge includes capacitance of all interconnecting cables and tonearm wiring to ground (or between the conductors), capacity added by the connectors and switches. Finally the internal wiring of the preamplifier circuit and preamplifier input circuit capacitance, which varies widely depending on the circuit design, adds capacitive loading to the cartridge, Fig. 27-27. In many cases the total capacitance that appeared as a capacitive load for the cartridge exceeded 1000 pF, which resulted in an electrical resonance peak around 7–8 kHz followed by premature response rolloff at frequencies above this point.

27.7 Phono Preamplifiers

Phonograph cartridges require a special type of amplification to reproduce the recorded sound the way it existed during the recording session. The electrical signals from the cartridge, measuring only a few millivolts rms have to be amplified into signals of many volts. This has to be accomplished with:

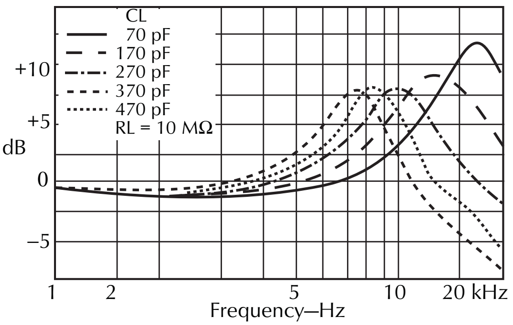
- Minimum distortion.
- Flat frequency response.
- Excellent SNR.

The phono preamplifier has to amplify a cartridge signal without:

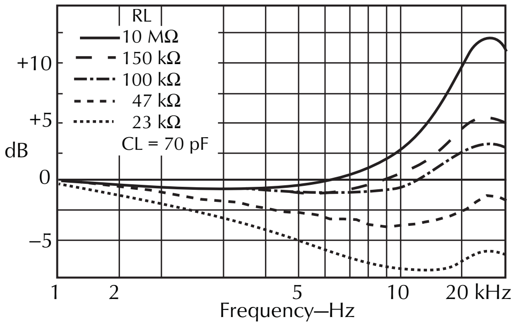
- Changing its phase.
- Adding more than a small percentage of harmonic and intermodulation distortion.
- Adding to the noise content of the original signal from the cartridge.
- Needing enough reserve power to handle any unusually high transient signals.

The average required voltage amplification is 40–50 dB and is dependent on the output of the cartridge. The dynamic cartridge produces 4–5 mV of output for the average recording signal. A preamplifier gain of 45 dB will boost the signal output to nearly 1 V, the level required to drive most power amplifiers to full output. The noise contribution of the cartridge and of the recording medium requires the preamplifier noise level to be at least 70 dB below the average input signal of 10 mV.

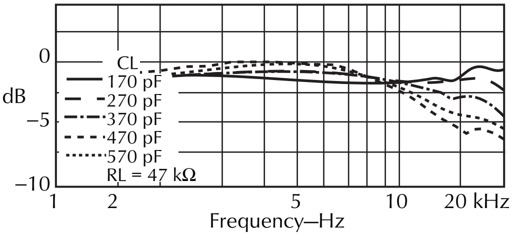
The frequency response of the circuit should follow the RIAA characteristics, with the low frequencies



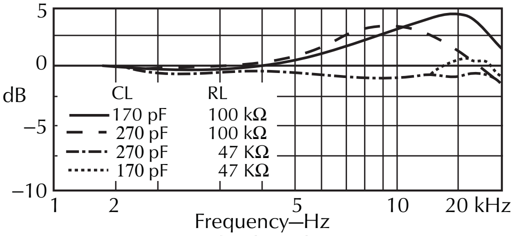
A. Effect of varying capacitive load on resistively unterminated cartridge.



B. Effect of varying resistive load on capacitively unterminated cartridge.



C. Effect of varying capacitive load on cartridge terminated with 47 kΩ.



D. Various load combinations.

Figure 27-27. Effect of cartridge loading on frequency response. Courtesy Stanton Magnetics, Inc.

boosted about 20 dB, and the high frequencies attenuated by the same amount, with respect to 1 kHz, which implies that the preamplifier with 40 dB of gain at 1 kHz will have as much as 60 dB of gain at 20 Hz and only 20 dB of gain at 20 kHz.

It is not unusual for a cartridge producing an output of several millivolts for the average modulation to produce 100 mV voltage peaks. Cartridges are designed to produce an output voltage of around 1 mV for each centimeter per second of recorded velocity or for the average recorded level of 5 cm/s, the cartridge output is 5 mV. Some preamplifier circuits when overloaded by fast spikes can recover in a matter of microseconds and resume their normal operation while others are incapable of recovering fast and once overloaded stay in this unbalanced state long enough to produce audible distortion of lower-level signals that may follow. Direct coupled stages, which don't employ large capacitors and inductors, have much higher slew rates and consequently react much faster and with less distortion to audio signals.

The average moving-coil cartridge produces from 0.1–0.6 mV output with the source impedance of a few ohms and an inductance of a few millihenries, so 20–30 dB of additional voltage gain is required from the pre-preamplifier. Because the output level of the cartridge is so low, an extra demand for low-noise performance is placed on the circuit. To maintain the same SNR as in high-output moving-magnet cartridges, the pre-preamplifier (or head-amplifier) circuit should have 20 dB lower noise than the preamplifier for the moving-magnet cartridges. One of the ways to achieve this lower noise is by using a step-up transformer. The power supply for the low-level amplifiers requires excellent regulation and extremely low ripple voltage.

The preamplifier input for the moving-magnet cartridges requires a 47 kΩ input resistance and a low, preferably adjustable, capacitive load. The proper termination of the moving-magnet cartridge is very important for the correct performance of the transducer. Moving-magnet cartridges have a resistive and inductive nature, so designers specify the capacitive load. If the specified capacitive load is higher than the total capacity of the circuit, the preamplifier should have a provision to add capacitance to the cartridge termination as required. If the total capacitance is larger than needed, cables can be made shorter or replaced with ones having lower capacitance.

27.8 Laser Turntable System

The Laser Turntable (LT) is manufactured by the ELP Corporation, Japan, Fig. 27-28. It features a contact-free optical pickup system allows records to be played thousands of times without damage to the record. The LT operates with five laser beams, three of which stabilize the groove and two that read the analog audio information from the record. The first two beams aim at the left and right shoulders of the groove for tracking. The next two read the stereo sound at 10 microns below the shoulder (the standard position). The final beam maintains the height between the laser head and the surface of record to manage thicker or warped records. The LT eliminates acoustic feedback and sound alteration and will play warped and rippled records (up to 5 mm deviation). For convenience, the LT comes with a remote control and can be paused and scanned much like a CD player.

Figure 27-28. ELP Laser Turntable. Courtesy ELP Corporation.

The same audio information on records is engraved from the shoulder to the bottom of a record groove. The laser reads audio information that is 10 microns below the shoulder, Fig. 27-29, therefore, the laser picks up audio information which has not been touched or damaged by a pickup. It plays the virgin audio information on the groove without digitization.

The incident area of the laser beam on the groove is one-fourth the contact area of the best stereo needle and twenty-six times smaller than a mono needle, Fig. 27-30. The laser beam travels to the wall of the groove and back. The reflection angle is transferred to the audio signal. Therefore, the LT maintains analog sound through the entire process, without any digitization. As a result, the LT cannot differentiate between an audio signal or dirt on the record, so the vinyl record must be absolutely clean and free of debris.

The laser beams must reflect from an opaque surface in order to be read. Clear or colored records are transparent, or translucent, and will not reflect light to the sensors. Other types of records that may have difficulty include:

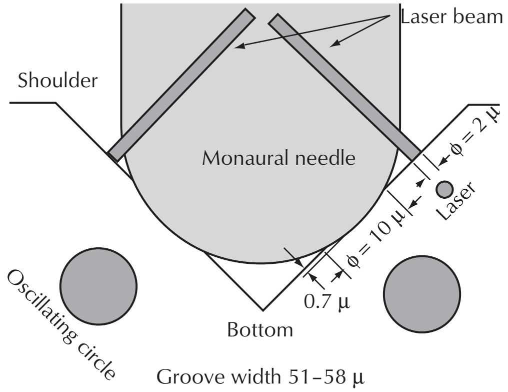

Laser beam
Shoulder
Monaural needle
ϕ = 2 μ
Laser
ϕ = 10 μ
0.7 μ
Oscillating circle
Bottom
Groove width 51–58 μ

Figure 27-29. The laser beam picks up the signal closer to the shoulder where the standard needle does not touch.

- Vertical cut records like the early Edison “Diamond Cut” series. The modulation is up and down rather than lateral.
- Rounded groove shoulder.
- A groove with a rounded bottom produces distortion.-

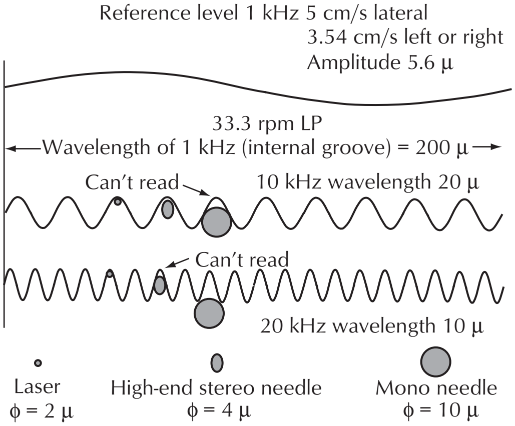

Reference level 1 kHz 5 cm/s lateral
3.54 cm/s left or right
Amplitude 5.6 μ
33.3 rpm LP
Wavelength of 1 kHz (internal groove) = 200 μ
Can't read
10 kHz wavelength 20 μ
Can't read
20 kHz wavelength 10 μ
Laser
ϕ = 2 μ
High-end stereo needle
ϕ = 4 μ
Mono needle
ϕ = 10 μ

Figure 27-30. The incident area of the laser beam is one-fourth the size of a stereo needle and twenty-six times smaller than a monaural needle.

No Acoustic Feedback or Sound Alteration. Feedback is typically caused by sound from the loudspeakers (or from elsewhere) reaching the turntable and mechanically picking up the vibrations, to be amplified again. There is no needle singing. The LP is safely in a drawer and the laser reads only the undulations of the groove, therefore there is no need for elaborate vibration isolation pads. The LT will not hear outside noises such as footsteps on the floor, door slamming, or other vibrations in the area.

Operating the Laser Turntable. When a record is placed in the drawer and the play button pressed, the record tray closes and the LT scans the disc to identify the various bands or cuts. The bands are displayed on the front panel Record Profile LCD display. A single vertical line above the bands indicator shows the position of the laser pickup head. The vertical indicator travels across the record as it is played, showing its exact position on the record, and which band is playing.

On the initial scan the laser head moves from the inside (spindle) to the outside track while marking the bands. The machine then moves into the first band and measures the distance from the head to the record surface. After a few seconds the record will begin playing from the beginning.

The LP can repeat the same record up to five times, repeat a cut, listen to a segment of the cut, play a single grove segment repeatedly, or play selected bands in any order.

When a record starts to play, the message window displays the rpm of the platter. When a record is playing, the display shows the elapsed running time, the elapsed time of the current cut, the remaining time of the side, and total time of the side.

27.9 Record Care Suggestions

One of the most effective ways to keep the sound from the record free of noise and unwanted pops and clicks is to keep the groove and the stylus clean. The causes for dirty records are obvious; accumulation of airborne dust, finger grease, cigarette smoke, and anything that can be attracted by the static charges that exist on the surface of the vinyl disc. The dirt around the playback stylus is mainly due to raking the groove. Dust particles, as they settle down on the record surface, are attracted by the stylus, especially if it has a static charge on it. Better cartridges have their styli electrically grounded to bleed any static potential from the cantilever assembly to ground.

27.9.1 Brushes

One method to keep room dust out of the record groove is to have the cartridge work with the dust-collecting brush. In sliding over the surface of the vinyl record, the electrically insulated brush produces a static charge of its own that attracts and holds the dust particles from the surrounding area. The stylus cantilever, which is metallic and electrically neutral because of grounding, stays clean and free to vibrate and track the modulation of the groove.

27.9.2 Record-Cleaning Machines

The groove modulations in vinyl LP's are so small, on the order of the wavelength of light, that any compound, be it liquid or solid, will cause distortion in the reproduction of those grooves. The diamond stylus can be equated to a rock, and the vinyl record to Jell-O. Picture a rock running through jell-o at a high velocity. Anything that changes the way this rock moves through the Jell-O will cause changes in the recorded sound.

In the groove is a conglomerate of fungus, mold, dirt, ash, pollution, mold release compounds, various cleaning fluids and preservatives, etc. All these substances affect the way the stylus reads the groove and will affect the sound. A good vacuum cleaning machine will allow you to scrub the record with cleaning solution and then vacuum the record surface clean of the fluid carrying the contaminates away with it. A record cleaned on a good vacuum cleaning machine is microscopically clean and will sound it.

One of the great shocks in audio is the first time you hear a record you know very well cleaned by a vacuum cleaning machine. The sound is cleaner, clearer, crisper, with the sound of the hall or acoustic space very easy to hear. A clean record will not wear out. It is not the stylus that ruins the records it's the stylus going through grunge and pressing the grunge into the vinyl grove that kills the sound of records.

Vacuum record cleaning machines all work the same way; a record is placed on a turntable, the turntable turns the record while the machine or the operator scrubs the record, the vacuum nozzle then sucks the contaminated fluid off the disc. A higher price gives you quieter operation or greater sophistication in application of cleaning fluids. In the end the result is pretty much the same. VPI's HW-16.5 (in production for almost 30 years), Fig. 27-31, is an inexpensive record cleaner. The VPI HW-27 Typhoon Record Cleaning Machine is twice as powerful as other cleaning machines. It includes a 7.9 A 120 Vac vacuum motor and an 18 rpm turntable motor, Fig. 27-32.

It is strongly advised that before using any cleaning device the instructions be followed precisely and some experimentation be done on a few records before the entire library is cleaned or covered with a preservative coating. A word of caution, if too much record preservative is used, it will do more harm than good. Not only does the excess of material not lower the surface noise,

Figure 27-31. VPI HW-16.5 basic record cleaner machine. Courtesy VPI Industries.

Figure 27-32. VPI HW-27 Typhoon Record Cleaning Machine. Courtesy VPI Industries.

but it contaminates the stylus tip to the extent that it is no longer able to stay in the groove. Accumulation of the cleaning or antistatic substance on the stylus tip also increases its dynamic tip mass, interfering with tracking of high-frequency modulation. Consequently, cleaning the cartridge stylus becomes as important if not more important than cleaning records.

27.9.3 Record Storage

The worst enemies of records are dust, heat, and mildew. To protect records from contamination they should be kept covered in their sleeves. Sleeves should be static free if possible. Records should be stored either vertically or horizontally (freshly pressed LPs are stacked one on top of each other to prevent warpage). If stacking horizontally, sizes should not be intermixed, and the stacks should be neat and not too high. If stored vertically the records should not be loose and should not be leaning; this will introduce warpage. Record cleaners or preservatives should not be applied prior to storage because there is a good chance of mildew forming on the records if they are stored damp.

27.9.4 Cleaning Records

Warning: Old 78 rpm records should never be washed with solutions containing alcohol or other chemicals that dissolve shellac, the major binding ingredient in the record material. Vinyl LP records are much more forgiving and can be cleaned with alcohol solvents. The safest and most effective cleaning solvents are simple household liquid soaps that can do the job well if certain precautions are followed.

Records should not be washed unless necessary. Dry clean them first with a soft brush or lint free velvet cloth. If the record must be washed, use distilled water; never use hot water or water containing dissolved minerals. Record labels should be protected by placing a piece of thin plastic over the labels. Use a soft camel hair brush or piece of moistened velvet with a couple of drops of liquid detergent or shampoo applied to clean the grooves in a circular motion. Rinse thoroughly with distilled water, and then wipe with a clean lint free cloth. The record can be blow dried with a hand dryer set to the cool position (never hot). Most of the dirt in the groove is dust attracted by the static charges that exist on the record surface. Washing or rinsing the record surface dissipates these electrical charges, allowing the dust to float away.

Turntable mats are the greatest contributors of dust contamination because turntables are left to stand open for prolonged periods of time, accumulating dust on the mat. When clean records are placed on the mat, the underside of the disc picks up most of the dust off the mat. It is important for the mat to be cleaned, even washed.

Vinyl records (and CDs) are sensitive to heat. When the record is pressed under very high pressure, vinyl is flattened into a thin plastic disc that is forced to cool down under pressure until the vinyl is no longer pliable. Then the disc is cooled down further to room temperature. The forces applied to the plastic during stamping remain in the record. If the record is exposed to elevated temperatures again, the forces retained within the material will be released and the disc will warp. Once this happens, the disc is destroyed. Leaving the disc in a closed car or on a window sill on a sunny day will accomplish this.

This page intentionally left blank

Chapter 28

Magnetic Recording and Playback

by Doug Jones and Dale Manquen

28.1 Introduction 1041
  28.1.1 The Family of Magnetic Recording Devices. 1042
  28.1.2 Tape Recorder as a Transformer 1042
  28.1.3 Changes with Time and Space 1042
28.2 Tape Transports 1043
  28.2.1 Tape Metering 1043
    28.2.1.1 Tape-to-Capstan Contact Enhancement 1046
    28.2.1.2 A Word of Caution Regarding Urethanes 1046
  28.2.2 Flutter 1047
  28.2.3 Tape Tensioning 1047
    28.2.3.1 Capstan-Derived Tensioning 1048
    28.2.3.2 Spooling-Motor-Derived Tensioning 1049
  28.2.4 Tape Guiding 1050
    28.2.4.1 Tape Height 1051
28.2.4.2 Head Azimuth 1051
    28.2.4.3 Tape Guides 1052
28.3 Magnetic Heads 1053
  28.3.1 Geometric Characteristics 1053
    28.3.1.1 Gap Length Loss 1053
  28.3.2 Spacing Losses and Thickness 1055
    28.3.2.1 Equalization Boosts 1056
    28.3.2.2 Fringing 1057
    28.3.2.3 Contour Effect 1057
    28.3.2.4 Crosstalk 1058
  28.3.3 Frequency Characteristics 1058
    28.3.3.1 Inductive Rise 1058
    28.3.3.2 Hysteresis Loss 1058
    28.3.3.3 Eddy Current Loss 1058
  28.3.4 Combined Characteristics 1059
  28.3.5 Noise 1059
    28.3.5.1 Track Width 1060
    28.3.5.2 Thermal Noise 1060
    28.3.5.3 Barkhausen Noise 1060
    28.3.5.4 Magnetostrictive Noise 1061
  28.3.6 Record Heads 1061
    28.3.6.1 Biased or Anhysteretic Recording 1062
  28.3.7 Erase Heads 1063
  28.3.8 Head Degaussing (Demagnetizing) 1064
  28.3.9 Tape Components 1065

28.3.9.1 Base Films 1065
28.3.9.2 Binders 1065
28.3.9.3 Magnetic Particles 1066
28.3.10 Magnetic Performance Curves 1067
28.3.11 Magnetic Tape Specifications. 1068
28.3.12 Problems with Older Tapes. 1071
28.3.12.1 Adhesion and Peeling Oxide 1071
28.3.12.2 Brittleness 1071
28.3.12.3 Splice Failures 1071
28.3.12.4 Print Through 1072
28.3.12.5 Sticky Shed and Tape Baking 1073
28.3.12.6 Squealing Tape 1074
28.4 Analog Circuits and Systems 1074
28.4.1 Playback Amplifiers 1074
28.4.2 Record Circuits 1076
28.4.3 Bias and Erase Circuits 1078
28.4.4 Noise Reduction Systems 1078
28.4.5 Sync Operation. 1079
28.5 Digital Magnetic Recording 1079
28.5.1 Longitudinal Digital Tape Transports. 1079
28.5.1.1 Signal Flow 1079
28.5.2 Helical Scan Digital Tape Transports 1080
28.5.3 Rotary Digital Audio Tape 1080
28.5.4 Packing Density Maximization with Rotary Head Recorders 1081
28.5.5 Heads for Digital Tape Recorders and Hard Disk Drives. 1082
28.5.6 Magnetic Disks 1083
28.5.6.1 Floppy Disks 1083
28.5.6.2 Rigid Disks 1084
28.5.7 Hard Disk Drives 1084
28.5.8 Hard Disk Electronics 1086
28.5.9 Formatting Media 1087
28.5.10 Long-Term Storage 1087
28.5.11 Data Interchange 1088
28.6 Tape Recorder Transport, Maintenance, and Testing 1088
28.6.1 Speed 1089
28.6.2 Flutter 1091
28.7 Tape Testing 1093
28.8 Magnetic Head Troubleshooting and Maintenance 1094
28.8.1 Head Relapping 1095
28.9 Routine Signal Alignment Procedure 1095
28.10 Automated Alignment 1097

28.1 Introduction

Many things have changed since the first, second, and third editions of this book were written. Magnetic recording is may still be the dominant storage technology, but its dominance is slipping. The everyday use of analog reel-to-reel recorders and longitudinal DASH digital tape recorders is virtually gone as are helical scan modular digital multitrack recorders. Computer-based systems storing data on random access hard disk drives are still popular, but as the price comes down on memory-based recorders that have no moving parts at all, these devices will certainly win out in the end.

The driving factor behind this shift is economy in both the acquisition and operating costs of the newer formats. The new systems utilize techniques and components that were developed and are mass produced for the consumer and computer markets, not just the very limited professional audio market.

We will first explore the underlying technologies common to all magnetic recorders, old and new. In spite of the rapid growth of digital techniques in audio, analog recording is by no means dead. Many albums are still recorded and mastered using analog audio recorders. In addition, the audio (and video) archives of the second half of the twentieth century are stored in vaults on millions of reels of analog tape. As a result, we will need qualified tape recorder operators and maintenance technicians for many years to come.

Unfortunately, however, much of the knowledge about analog audio recorders and the mentors who taught this information are slipping away. This chapter will provide an overview and in some cases more detail than the casual reader requires. A full treatment of the theory and practice of analog recording and specifically magnetic recording is clearly beyond the scope of this work. We hope that this will whet the appetite of some and spur others on to write more comprehensive treatments.

The roots of the modern-day tape recorder can be traced to Germany in the mid-1930s. Two German companies, AEG and IG Faben, worked together to develop the concept of recording on a coated tape. AEG built the machine and IG Faben developed the magnetic particles and manufacturing methods for the tape. Although few people outside Germany took note, the German tape recording industry flourished, with 5 million meters of tape being produced in 1939, Fig. 28-1. The German machines, using a form of plastic tape, were vastly superior to English recorders using steel tapes and American recorders using spools of steel wire.

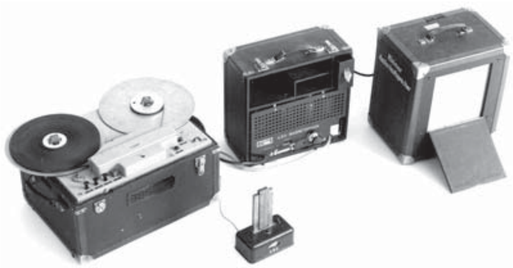


Figure 28-1. German Magnetophon, 1935. This portable magnetophone weighed over 100 pounds. Courtesy BASF Corporate Archives.

At the end of World War II, the victorious Allies set aside all of Germany's patents as a form of reparation for the war. As a result, the wealth of German tape recorder technology was quickly and freely exploited in the United States and many other countries. Within just a few years tape recording replaced disk cutting as the primary method of recording information.

Magnetic tape offered several valuable advantages over phonograph disk recording and even the American fledgling wire recorders, including improved signal to noise ratio, lower distortion, and better frequency response. None of these quality features triggered the quick adoption of tape recording by the American radio networks.

The networks needed the ease of use, especially the ability to create undetectable edits by cutting and splicing the tape. The recorders also had an important secondary benefit—the ability to stop and quickly restart recording. (One does not stop a record cutting lathe in the middle of a cut!)

Both of these features stem from the nature of the tape recording. Tape recording is a serial process that distributes the audio events on a very long piece of tape. The time of the event is implicitly encoded in the position of the event along the tape. The editing scissors and tape now become a *time machine* that can alter the apparent time of an event by relocating the tape segment of the event to a new position in the reel. Editing is merely playing tricks with this *time machine* to remove or replace events to alter the program.

Ironically, more than 50 years later the pendulum has swung back, turning the serial nature of tape into a shortcoming. Most operations in a recording studio require one or more replays of previously recorded material. During a mixdown session, for example, the same song may be replayed 500 times before the final mix is finished. Each of these replays requires time to rewind to the head end of the desired selection. A 3-minute tune recorded at 30 in/s occupies 450 ft of tape. If the tape recorder takes 15 seconds to rewind the tape, the engineer would spend 500 15 second intervals waiting for the tape recorder to rewind. That is over 2 hours spent in rewind mode!

Contrast this sluggish operation with a digital audio system's hard disk that can locate any position on the disk in less than 10 ms. 500 rewinds might now take less than 5 s!

28.1.1 The Family of Magnetic Recording Devices

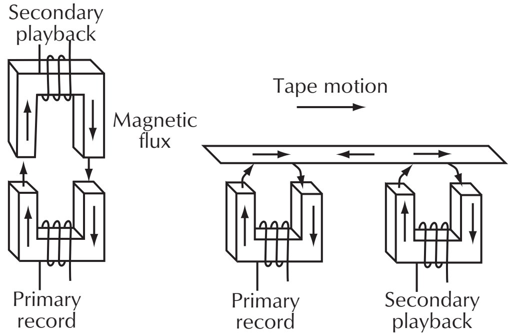
All magnetic tape recorders are members of a larger family of storage devices that utilize moving storage media. Other members of this family include phonograph disk recorders, motion picture cameras and projectors, optical laser disks for video and audio, and magnetic disk devices for computer data storage. These storage devices share one very important characteristic—they all are complex electromechanical devices. In addition to electronic circuits that amplify, process, and control the basic signal that is to be recorded and retrieved, each device also contains numerous mechanical devices to move the media past the recording and reproducing transducers and also position the transducers for optimum performance.

All magnetic recorders share key features:

1. The recording process is instantaneous, requiring no intermediate processing before the signal can be replayed.
2. The record and playback processes exhibit *reciprocity*, meaning a single transducer may be used for recording or playback.
3. The storage medium can be easily erased and reused.
4. The parameters of the system (speed, track width, encoding scheme, etc.) can be customized for a broad range of audio and video applications.

28.1.2 Tape Recorder as a Transformer

A magnetic tape recorder can be visualized as a specialized form of transformer. In a conventional transformer, an electrical signal on the input or primary winding is converted to magnetic energy in the magnetic core of the transformer. This magnetic energy is then converted to an electrical signal in the output or secondary winding, proportional to the ratio of the windings. Transformers can be quite efficient and the losses are typically just a few percent of the total power passing through the device. The best audio transformers introduce only very small amounts of distortion to the amplitude and frequency response of signals passing through the transformer, Fig. 28-2.



A. Transformer. B. Tape recorder.

Figure 28-2. Transformer versus tape recording.

For a tape recorder, the input and output windings consist of the record head and reproduce head. The *magnetic core* that couples these windings is a conveyor belt covered with magnetic particles in the form of the magnetic tape. A magnetic image is permanently impressed on the conveyor at the record head. When this image passes over the reproduce head, anywhere from milliseconds to years later, the magnetic image creates a signal in the head that is analogous to the original signal.

Unlike the fixed transformer core, the recording tape is fraught with numerous distortions, losses and imperfections that require attention. Virtually every component in the record/reproduce chain, including the heads, tape, signal electronics circuitry, and mechanical drive system, contributes to these errors.

28.1.3 Changes with Time and Space

Those of us who work in audio and acoustics are used to thinking of audio as a complex wave made up of discrete frequencies. These frequencies move through the medium (air) at a fairly constant speed. Distortions due to changes in propagation time are not often observed. In the equation

$$\lambda = \frac{C}{f} \tag{28-1}$$

where,

C is the speed of sound in air and is essentially viewed as a constant.

In the analog tape recorder, C, the speed of the tape past the tape head, is a variable. Changes in the speed of the tape past the record/play heads will almost always be noticeable and, unless you are trying to sound like a small rodent, undesirable! The mechanism responsible for moving the tape past the heads in a constant and repeatable manner is called the *transport*.

28.2 Tape Transports

The beginnings of modern-day tape transports can be traced to Vlademar Poulsen, the Danish inventor of the magnetic wire recorder. Poulsen's experiments in 1898 consisted of moving an electromagnet along a piece of steel wire to record and reproduce sound. He soon learned, just as every tape recorder operator today learns, that the relative motion between the transducer (the electromagnet) and the storage medium (the wire) must be uniform and repeatable.

Many of Poulson's solutions to this problem, such as sliding the electromagnet down a long, sloping wire worked reasonably well but were hardly practical! The functions of his transport device, however, were the same as modern tape recorders, specifically:

1. To drive the tape (or wire) at a repeatable, and preferably constant speed over the surface of the transducer heads.
2. To maintain a fixed mechanical alignment of the tape as it crosses the heads.
3. To provide contact pressure between the tape and head by either tensioning the tape or pushing the tape against the head.
4. To provide the necessary auxiliary motions of the tape required for functions such as rewind, search, and editing.

The early German Magnetophon developed by I.G. Faben in the 1930s satisfied all of these requirements with a simple mechanical layout. Over 70 years later, today's recorders have essentially the same layout, shown in Fig. 28-3. The reels of tape are mounted on the shafts of two motors that provide the high-speed spooling and the play-mode tape tensioning. The tape moves from the supply reel on the left to the takeup reel on the right. As the tape leaves the supply reel, it is steered by guides to pass over the erase, record, and playback heads. Following the heads is a constant-speed tape drive consisting of a rotating shaft called a capstan and a pinch roller to press the tape against the surface of the capstan. The tape then passes to the takeup reel, Fig. 28-3.

This layout was used on virtually every tape recorder ever built except for the infamous Ampex 400 built sometime in the early 1950s that placed the capstan/pinch roller assembly to the left of the heads.

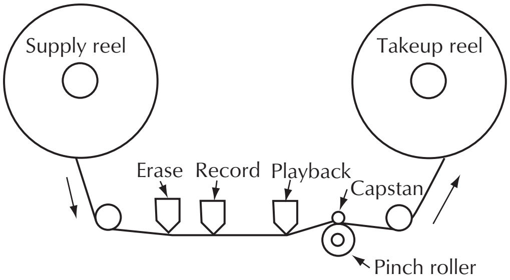


Figure 28-3. Classic tape transport layout.

The typical degree of precision that is available today in a professional recorder includes a tape speed variation of a few hundredths of a percent, mechanical alignments of less than one-thousandth of an inch (0.001 inch) and three-thousandths of a degree (0.003°), and tension variations of a few percent. Even these seemingly small variations create readily observable errors in recordings, leaving opportunity for future improvements.

28.2.1 Tape Metering

Ever since the early introduction of tape recorders to radio broadcasting, it was desired to have world standards that would permit tapes to be freely exchanged between facilities around the world. Furthermore, it was necessary to be able to freely exchange segments within a reel by editing. This requires absolute speed accuracy throughout the reel.

Broadcasters were concerned about the running time of a radio show. If the show was timed at exactly 30 min when it was recorded, it should also play in exactly 30 min on the air. A common timing accuracy specification of 0.2% means that the tape could play up to 0.2% of 30 min fast or slow or 3.6 s of error in either direction. This could result in either 3.6 s of overlap with the subsequent program, or 3.6 s of dead air silence while waiting for the next show to start.

A more demanding speed specification is the absolute speed error throughout the reel. If the tape machine runs 1% fast at the beginning of the reel, and 1% slow at the end of the reel, the overall timing might come out just fine. But when you cut a segment of music from the head of the reel into a song at the end of the reel, you now have a 2% speed jump at the splice, with a very noticeable pitch change.

A simple speed control technique is to clamp the tape to a surface that is moving at the desired tape speed, such as the outer periphery of a rotating drum. The tape is thus forced to move at exactly the same correct speed. Various implementations use drums that range from over 2 inches in diameter to tiny shafts less than 0.1 inch in diameter. In general, the larger the drum, the more accurate the tape speed control. The very small spindles are usually employed at the very slow tape speed found with compact cassettes and consumer videocassettes.

The rotating drum is called a capstan, named after a device used on sailing ships to pull in cables and hawsers, and the clamping device is called a *pinch roller*. The simplest capstan is the shaft at the end of a motor. The diameter of the shaft is chosen so that the shaft's circumference will move at the desired linear tape velocity when the motor is spinning at operating speed.

The actual linear velocity of the capstan surface is slightly lower than the tape's speed. The effective speed of the tape is measured at the *neutral axis* of the tape, about ½ of the tape thickness into the tape for large capstans, but dropping down to about ⅓ of the tape thickness for small capstans. Remember that the total thickness of a tape is the sum of the backing and coating thicknesses. A nominal 1.5 mil tape is really about 2 mils thick—1.5 mils of backing substrate and 0.5 mils of oxide coating.

Other designs use a capstan/flywheel assembly that is driven by belts or rubber-tired idlers that engage the primary drive motor's shaft. The resulting reduction in rotational speed permits the use of a larger capstan diameter. A good example with a belt reduction is the 3M Isoloop™ tape transport. At 15 in/s the capstan motor spins at 30 rev/s, but the large capstan turns only 2½ rev/s. The 12:1 speed reduction permits large diameters on all drive surfaces.

A flywheel in normally employed on the capstan's shaft to smooth out any small speed variations. The effectiveness of the flywheel increases directly with increased flywheel moment of inertia, but inversely with the square of the diameter of the capstan. The large capstan diameter of the 3M transport required a flywheel weighing 6 pounds!

Any rotational speed disturbances in the capstan will show up as linear speed variation in the recording tape. This means that the capstan must spin at an absolutely constant speed. The simplest constant speed device is a hysteresis synchronous motor. *Synchronous* indicates that the motor runs at a speed that is locked to the frequency of the voltage driving the motor, similar to a clock motor (before battery operated clocks). The motor contains a pair of windings for each operating speed. The two windings are physically offset by ¼ of the distance the motor rotates during one cycle of the drive voltage. One of the windings in each pair is connected directly to the power source. The second winding is connected with a large capacitor in series with the winding to shift the phase of the current and the resulting magnetic field in the second winding approximately 90° with respect to the main winding. The physical and electrical shifts work together to create a rotating magnetic field.

Two-speed hysteresis synchronous motors are common in tape recorders, and a few three-speed motors are also used. The motor must be designed for the intended operating frequency of 50 Hz or 60 Hz and the phase-shifting capacitor chosen for the appropriate frequency.

Although the hysteresis synchronous motor is an economical solution, it has major shortcomings. First, the speed of the motor is only as good as the stability of the frequency driving the motor. We assume that ac source is stable. However the power companies only guarantee a certain number of cycles each day. At any given time the frequency of the grid at any location may be slightly high or low depending how much adjustment is needed to bring the daily total of cycles into compliance. Sometimes, however, it is desirable to run at other than nominal speed for special effects or pitch correction. This Variable Speed Oscillator (VSO) operation requires a versatile power source for the capstan motor that can be shifted in frequency. If the frequency is shifted very far from the nominal frequency, the phase-shift capacitor will no longer provide a true 90° of phase shift. The motor will begin to vibrate, the power of the motor will decline, and the motor's temperature may rise. The maximum practical speed shift is then less than 15%.

A third problem is that the selection of speeds is quite limited. For 60 Hz operation, there can be motor speed pairs of 3600/1800, 1800/900, 1200/600, and 900/450 rpm. If the shaft of the capstan motor is used as the actual drive surface, the desired tape speeds will determine the diameter of the shaft. Slow tape speeds require very small capstan diameters. The resulting small contact area can create speed errors due to slippage.

All these problems can be avoided by substituting a servo-controlled motor for the hysteresis synchronous motor. A servo-controlled motor utilizes a speed-measuring device on the capstan in the form of a high-resolution optical or magnetic tachometer. This tachometer may provide as many as 1200 speed samples per revolution of the motor, a rate high enough to detect not only overall average speed, but even very small speed transients due to imperfections in other components in the tape path. By comparing the speed sensed by the tachometer to a high-accuracy reference derived from a crystal oscillator, any variations or errors in speed are immediately detected. The control circuits use this error to generate corrections in the voltage driving the motor to cancel the speed error. The overall accuracy of this closed-loop system is primarily dependent on the accuracy of the tachometer and the reference clock.

The block diagram of a typical capstan speed control is shown in Fig. 28-4. Commonly referred to as a *phase-lock servo*, the system is, in essence, a clocked position detector. The servo automatically adjusts the motor voltage so that the tachometer will produce one pulse for each pulse of the reference clock.

The crystal oscillator/counter provides a highly accurate clock reference by dividing the frequency of the crystal oscillator down to a convenient lower frequency. The switching transition of the clock serves as a strobe to sample the position of the tachometer. If the motor is running exactly at the desired speed, each tachometer transition will coincide exactly with a clock transition. The phase comparator compares the tachometer and clock signals to determine which signal arrives first and the amount of timing error between the sources. The error signal generated by the phase comparator is amplified and passed through a low-pass filter that smooths the individual pulses into an average dc voltage that can drive a dc motor. This smoothed voltage is applied to a power amplifier, called a motor drive amplifier (MDA), that drives the motor.

The phase-lock servo permits convenient speed control at multiple tape speeds by selecting various points along the divider chain. The minimum speed is limited by the data rate from the tachometer and the smoothing provided by the low-pass filter. The maximum voltage and current available from the MDA typically sets the maximum speed. Speed ranges of 2:1, 4:1, and 8:1 are common in audio recorders with servo-controlled capstans.

Variable-speed operation for a servo system is much simpler than for the hysteresis synchronous motor. A simple variable-frequency oscillator can be substituted for the fixed reference to provide infinitely variable speeds!

The *de facto* standard for professional machines is that an external VSO frequency of 9600 Hz from accessories will drive a servo at nominal speed. This 9600 Hz signal can be substituted for the crystal's countdown signal at an appropriate point in the countdown chain before the final speed-determining dividers. The VSO signal is thus able to control the machine at any of the machine's running speeds.

If the tachometer is accurately mounted, if the tachometer samples occur frequently enough to provide precise sensing, if the control circuit sends the correction signal to the motor quickly so that errors are sensed as they start, and if the motor can respond swiftly to corrections in its control voltage, then the motor will turn at a constant speed. The string of "ifs" in the previous sentence is a clue to the complexity of this servo design. The results, however, of a good design are very impressive, with professional recorders being able to suppress mechanically induced speed variations to below 0.05% rms at 15 in/s (38 cm/s) on a routine basis.

The speed-sensing device need not be attached to the driving capstan for phase-lock operation. The tachometer can be mounted on a free-running idler that is driven by the tape, but the extra time delay introduced into the error signal renders the system more difficult to control. This delay usually requires a reduction in the

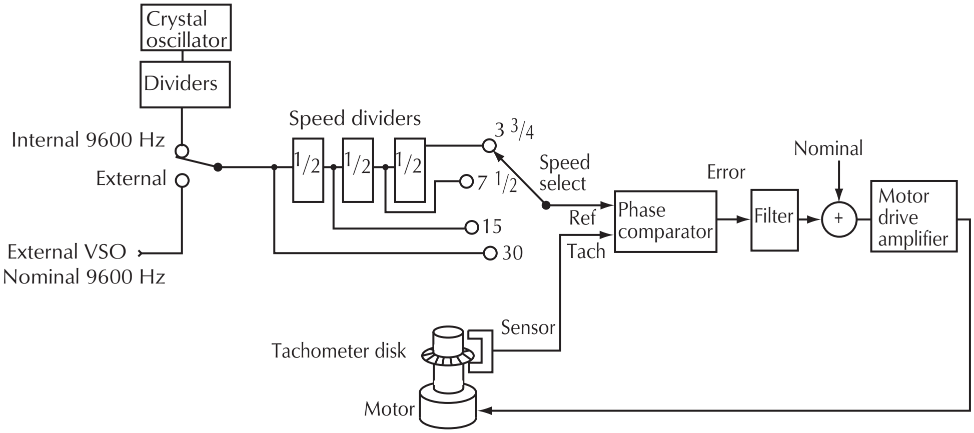


Figure 28-4. Capstan speed control block diagram.

stiffness and bandwidth of the servo loop, requiring either an improvement in the inherent errors that are to be corrected by the servo or a decrease in the expected level of performance.

A further extension of the free-running idler concept is to eliminate the drive capstan completely, as in machines manufactured by John Stephens. These machines relied on two high-performance spooling motors to perform all the speed control tasks. The Stephens recorders provided excellent speed control under normal conditions, but the large inertia of a 6 pound roll of 2 inch tape limited the responsiveness of such systems, rendering them vulnerable to abrupt disturbances such as tape splices or layer-to-layer adhesion of the tape.

28.2.1.1 Tape-to-Capstan Contact Enhancement

Constant tape speed requires that the capstan driving the tape must have enough traction on the tape due to friction to exert positive control of the tape. If we just wrap the tape around the capstan, the traction force due to friction will usually be too weak to exert full control of the tape. To maintain control, the capstan's drive force must be at least equal to the difference in the tape tensions on the ingoing and outgoing side of the capstan, as shown in Fig. 28-5.

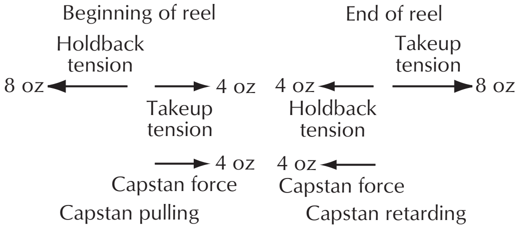


Figure 28-5. Forces at the capstan.

Active contact enhancement devices such as the rubber pinch roller push the tape against the capstan surface to maintain firm contact. Unfortunately, the pinch roller also produces numerous undesirable side effects, including:

1. Heavy side loads on the capstan that produce bearing wear and can even cause small diameter capstans to bend or tilt.
2. Speed errors due to the elastic deformation of the rubber roller at the point of contact.
3. Increased variations in speed created by imperfections of the rubber, eccentricities of the roller, and bearing rattle.

One way to avoid the problems of pinch rollers is to clamp the tape to the capstan using air pressure and a vacuum pump. Computer tape drives have frequently used hollow vacuum capstans to achieve rapid tape start/stop and shuttling. The capstan must be of porous material or have machined passageways so that air can be sucked from the surface of the capstan. The ambient air pressure will then push the tape firmly against the surface of the capstan. Since the air pressure differential will be somewhat lower than the maximum 14.7 pounds per square inch (psi) of nominal atmospheric pressure, there will need to be a substantial tape contact area to generate the required traction force.

Passive contact enhancement methods concentrate on maximizing the traction between the tape and capstan surface. Roughening of the capstan surface by sandblasting or coating the surface with urethane rubber or diamond-impregnated grit yields an improvement in the coefficient of friction. After heavy usage, however, the roughening will be polished away by the abrasive surface of the tape, or the urethane surface will glaze and harden, requiring reconditioning to avoid slippage.

Other passive techniques concentrate on eliminating any loss of contact due to air being trapped between the tape and capstan. This *air bearing effect*, which becomes evident at tape speeds as low as 30 in/s (78 cm/s), can be minimized by cutting bleed slots in the surface of the capstan. These slots are similar to the tread grooves on an automobile tire, providing escape paths for the trapped air.

28.2.1.2 A Word of Caution Regarding Urethanes

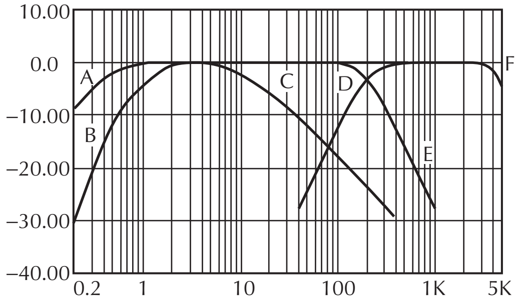
The standard roller rubber is neoprene, a fairly stable rubber compound that can resist ozone and smog. Many newer compounds, especially various urethanes, have also been tried with some success. Sometimes the new roller will give excellent results when new, but then it will glaze over and lose its adhesion to the tape. In other cases the roller's elastomer will turn into a gummy ooze with the consistency of taffy.

The urethane is affected by temperature and humidity conditions, and by any solvents used to clean the tape path. Always check the cleaning pad after you clean the pinch roller. If the pad has just tape residue, you are providing proper cleaning. If, on the other hand, you see a residue that looks suspiciously like the surface of the roller, you may be dissolving your pinch roller!

28.2.2 Flutter

Regardless of the passive and/or active contact enhancements, servo design, and workmanship standards employed in a given transport, some residual amount of tape speed variation will still be present. The long-term or fixed component of this speed error is denoted as speed accuracy, timing accuracy, or drift. The small, rapid changes in instantaneous speed are referred to as *flutter*.

Flutter is further broken down into three frequency bands, based on perceptibility, Fig. 28-6.



"B-C" is weighted, "A–E" is unweighted, "D-F" is scrape, and "A–F" is wideband. (Ideal response data graphed using Audio Precision System One Software.)

Figure 28-6. Flutter bandwidths.

Speed variations at rates up to a few cycles per second are termed *wow*, with the listener perceiving a cyclic pitch variation in music. The most common source of wow is eccentric rotating parts.

Faster flutter rates due to motor torque pulsations and rattling bearings add a fluttering sound to the music. As the flutter rate increases beyond a few hundred hertz, the listener no longer distinguishes the flutter components from the music. Instead, the listener notices a loss of crispness and clarity, with high frequencies created by percussion, strings, and brass sounding dull or mushy. These high-frequency scrape flutter components are generated as the surface of the tape scrapes over stationary elements such as fixed guides and heads, creating vibrations in the tape similar to the plucking of a stringed instrument.

Historically, wow and mechanical flutter have received much more attention than scrape flutter. In fact, tape recorders were used for music recording for nearly 20 years before the first transport with low scrape flutter was introduced. Even today designers of both transports and tapes treat scrape flutter more as an afterthought than as a primary problem, failing to quote any specifications for scrape flutter performance. Unfortunately for the user, the subjective evaluation of the clarity of a recording is very dependent on the amount of flutter in all three flutter bands.

Weighted peak flutter is an attempt to characterize a human listener's perception of flutter. Many years ago, numerous tests showed that the test subjects most readily identified flutter disturbances that occur at a rate of approximately 4 Hz. Furthermore, the tests indicated that the listener responded to the peak levels of flutter, even though the peaks may have been infrequent. Based on these test results, flutter meters now include bandpass filters peaked at 4 Hz and quasi-peak metering.

Today, every professional tape recorder produced in the past 35 years includes components to reduce scrape flutter, but the typical weighted peak flutter meter is totally incapable of measuring these components to verify proper performance!

In addition, misbehaving servo-controlled transports can generate flutter frequencies at virtually any frequency. Unlike the older machines with all their mechanical resonances below 100 Hz, newer machines can have servo oscillations well beyond 1 kHz.

The entire flutter spectrum should be measured, especially when performing maintenance testing of professional audio recorders.

28.2.3 Tape Tensioning

Magnetic recording tape, like all elastic media, must be stretched slightly to produce tension within the tape. For normal recording applications, the tape is stretched approximately 0.1% to achieve a typical tension of 4 oz per ¼ inch of tape width. Since this small amount of stretch is less than one tenth the level of stress required to permanently deform the tape, no permanent deformation results.

Four separate and often conflicting functions are performed by tape tension on a tape recorder:

1. Tape tension holds the moving tape firmly against the record and playback heads to achieve good high-frequency performance.
2. Tension stiffens the tape on the tape guides so that the tape position will remain constant.
3. Tension controls the stacking of the layers of tape on the takeup reel.
4. On machines without pinch rollers, the tension holds the tape against the capstan to create enough drive traction for proper tape speed control.

The classic tape transport of Fig. 28-3 utilizes the supply reel spooling motor to generate tape tension over the heads in the Play mode. The supply motor is energized in the clockwise (rewind) direction with a reduced voltage, generating a constant torque from the motor. To convert motor torque to tape tension, divide the torque by the radius of the tape pack (the lever arm).

However, the radius of the pack on the supply reel decreases as the tape plays off. By the end of a 10½ inch NAB reel, the radius has dropped to half the starting value, causing the tape tension to double. (Some plastic 7 inch reels have outside to inside diameter ratios of more than 3:1.)

The tape tension is further altered to some degree by every component that comes into contact with the tape. When tape slides over any stationary guide or head surface, the tape tension changes slightly due to the friction between the tape and the stationary surface. (The bearing friction and viscous drag of rotating guides is usually negligible.) The relative contribution of friction tension to the total tape tension ranges from a low of 5% for transports with only rotating guides to over 50% for transports with numerous fixed guides and/or large tape deflection angles around fixed guides.

The amount of drag tension generated by a cylindrical post is shown in Fig. 28-7. The tension and friction build up as the tape moves around the guide. The true expression for the total drag is an exponential function, but for tape paths with only small amounts of wrap, we can approximate the tension change with the expression:

$$\text{Tension change} = K \times \text{tape tension} \times \text{angle of wrap} \times \text{coefficient of friction} \qquad (28\text{-}2)$$

Note that although the diameter of the guide does not appear in the tension expression, the pressure exerted by the guide against the tape surface increases as the diameter decreases. This increased pressure makes small guides wear faster and accumulate dirt more quickly. Since a speck of dirt trapped on the surface of a small guide would also be more prone to scratch the tape surface, small-radius fixed guides must be kept very clean.

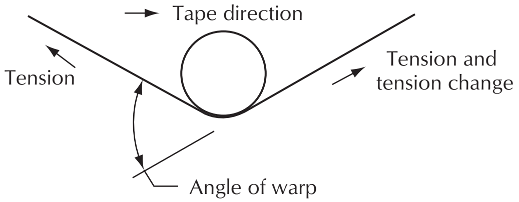


Figure 28-7. Tension increase due to guide friction.

The coefficient of friction depends not only on the type of tape, but also the condition of the roll of tape. Older tapes may lose the surface lubricants that allow the tape to slide freely across the stationary surfaces. This may result in a squealing sound as the tape runs through the recorder. Even worse, when sticky shed debris from breakdown of the urethane binders in the tape collects on the tape guides, the tape may be dragged to a dead stop. These problems occur commonly when dealing with older archived tapes.

Some transport designs are more sensitive than others to changes in tape tension resulting from tape problems. Tape paths with either high amounts of drag tension or no pinch rollers may require a readjustment of the tape tension to maintain acceptable performance and avoid tape slippage.

28.2.3.1 Capstan-Derived Tensioning

Because of conflicting tension requirements at various point along the tape path, it is useful to break the total path into segments with different tape tensions. The classic tape transport, for example, has two distinct *tension zones*, one to the left of the capstan and another to the right of the capstan. One common approach is to use a driven capstan and pinch roller as an isolation device, with the capstan motor supplying the power needed to overcome the tension differential across the capstan, see Fig. 28-5. This isolation can then be used, for example, to achieve low head tension and high takeup spooling tension.

Carrying this strategy one step further, if one capstan can be used to isolate the head from the takeup system, then it should be possible to use a second capstan to isolate the heads from supply reel disturbances. This would provide the desired isolation on both sides as shown in Fig. 28-8. The difficulty is in generating a controlled tension between the capstans where the tape passes over the heads.

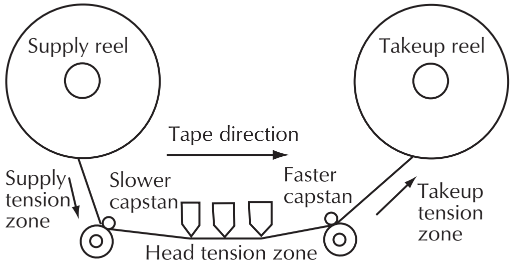


Figure 28-8. Dual capstan transport showing multiple discrete tension zones.

The solution is to have slightly different surface velocities on the two capstans. If we need a 0.1% stretch of the tape to give us the desired 4 ounces of tape tension, then the outgoing capstan must have a surface velocity 0.1% higher than the incoming capstan. This can be achieved by using two hysteresis synchronous motors with slightly different capstan diameters. The Gauss high-speed tape duplicators used this technique with great success. A very similar technique is to use a nonstretching plastic belt to couple the drive motor to both capstans. If both capstan shafts are identical, but the pulley on the outgoing capstan is 0.1% smaller than the other pulley, the desired speed differential will be realized.

Many dual capstan cassette decks provide bidirectional operation. A simple trick is to use an elastic rubber belt to drive both capstans, Fig. 28-9. Since the belt is elastic, it will stretch slightly whenever it delivers a pulling force to a load. It must pull on the incoming capstan's pulley with sufficient force to overcome the holdback tension, the friction due to the incoming pinch roller and capstan's bearings, and the load caused by the elastic deformation of the pinch roller at the point of contact with the tape and capstan. As a result, the stretched belt leaving the incoming pulley will have a slightly higher linear velocity due to the stretching. As this stretched belt passes around the outgoing pulley, the higher linear velocity will turn the outgoing pulley slightly faster than the incoming pulley. The difference in speed generates the desired tape tension. Everything is symmetric, so if the motor is reversed, both the tape and the tape tensioning will reverse.

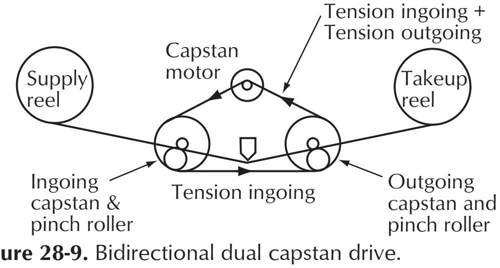


Figure 28-9. Bidirectional dual capstan drive.

These dual-motor and single-motor designs are both classified as tight loop or closed loop tape drives. Closed tape loops were first developed for the very high-performance recorders used to gather telemetry information from rocket testing. If the capstans are free of flutter, these designs can yield low mechanical flutter and wow since the heads are isolated from spooling disturbances. As a bonus, a closed loop tape path yields low-scrape flutter because only a short span of tape is free to vibrate at the heads.

Yet another closed loop design, the 3M Isoloop™, achieves the effect of two capstan diameters by using a single capstan with multiple alternating rings of large and small diameters. The step between rings is so small, on the order of 0.1%, that specially contoured pinch rollers can press the tape against the smaller-diameter rings on the incoming side and against the larger-diameter rings on the outgoing side of the capstan, Fig. 28-10.

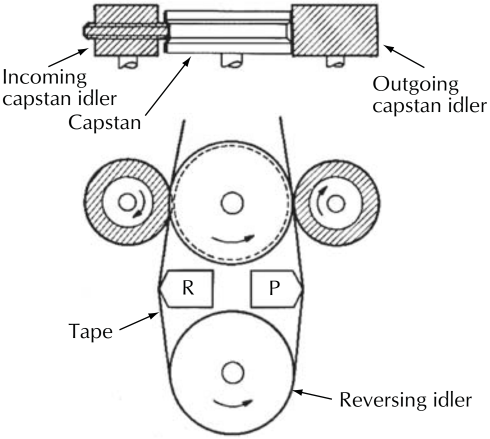


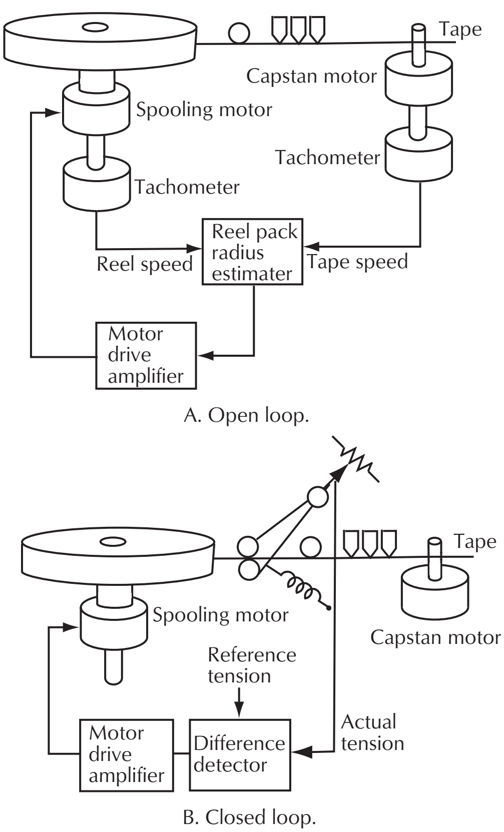
Figure 28-10. 3M Isoloop™ drive.

Unlike recorders that derive tape tension by controlling torque on the spooling motors, the tension of the closed loop drives varies slightly with tape thickness. Since the change in tape length is always constant, lower tensions are generated in thin tapes that stretch more easily. This decrease in tension is generally unnoticed since the thinner tape conforms more readily to the face of the heads, offsetting any pressure reduction.

28.2.3.2 Spooling-Motor-Derived Tensioning

The classic tape transport of Fig. 28-3 experiences a 2:1 tension change from beginning to end of reel. For nearly 25 years the recording industry was forced to struggle with recorders that had this doubling of tape tension, with attendant speed variations, splicing problems, and tape guiding variations. The advent of economical integrated circuits has led to more sophisticated designs that replace the constant holdback torque with an active tension servo control.

Tension servos fall into two categories—closed loop and open loop, Fig. 28-11. The closed loop servos directly sense the tape tension with a spring-loaded surface in contact with the tape. The tape pushes against the spring, causing a displacement of the sensing device. The resulting output voltage or current provides an electrical signal proportional to tension that can be used to control the spooling motor drive voltage.



A. Open loop.

B. Closed loop.

Figure 28-11. Closed and open loop tension servos.

Numerous types of sensing devices have been employed by various manufacturers; including photocells, photo-potentiometers, rotary potentiometers, and Hall effect devices. Noncontacting photosensors and Hall devices have demonstrated longer lives than rotary potentiometers with sliding mechanical contacts.

Rather than directly measuring the tape tension, the open loop tension servos sense other parameters such as the rotation rate of the spooling motors and infer or calculate the amount of drive to the spooling motor necessary to achieve the desired tape tension. For example, the rotational velocities of the spooling motors can be measured with dc tachometers attached to the spooling motor shafts. The tape speed can be derived from the frequency of the tachometer pulses coming from the capstan servo motor. Dividing the tape speed by the reel rotation rate yields a value proportional to the tape pack radius. Tape tension times tape pack radius is equal to the required motor torque. Therefore,

$$V_{motor} = tension\ adjustment\ factor \times \frac{speed_{tape}}{speed_{rotational}} \tag{28-3}$$

If the calculation is executed with analog circuits, the multiplication is easily implemented by a potentiometer, but the division requires an analog multiplier/divider integrated circuit.

The same results can be achieved with a read-only memory lookup table that is programmed with the correct motor voltages for various combinations of speed and pack diameter.

MCI tape transports using the open loop tension control method described above provided very accurate tension control over a wide range of speeds and reel diameters. The major shortcoming is that the calculated method cannot detect tension abnormalities due to bent reels, motor problems, or changes in friction.

A further benefit of the diameter calculations is the ability to anticipate the end of a reel of tape. Both unintentional unthreading of the machine and abusive high-speed unthreading can thus be avoided.

The MCI tension control worked very well because an integrated circuit performed the analog division with high precision. Other tension controls, such as the Tentrol for Ampex transports, substituted a pair of adjustment potentiometers for the analog divider. The resulting straight-line approximation of the division process was not as accurate as the MCI method, but it was much better than the 2:1 tension change without any sensing and control.

28.2.4 Tape Guiding

For proper recording and playback of a magnetic recording to occur, the tape must move over the heads in a very precise path. This tape path should be the natural path that the tape would follow without any external vertical constraints. The purpose of the guiding system is not only to protect the tape and to overcome the slight reel-to-reel variations in tape such as twists and bends due to tape-manufacturing tolerances but not to force the tape to perform any unnatural acts. Any such use of brute force will lead to tape damage, excessive guide wear, and/or instabilities and jumping of the tape.

The tape guiding system deals with five aspects of the tape motion—height, azimuth, zenith, wrap, and rack—with primary concern for the motion of the tape at the heads. Each aspect is in turn composed of two components: fixed errors due to misadjustment and dynamic errors due to tolerances and tape variations.

28.2.4.1 Tape Height

Height must be controlled so that the recorded tracks on the tape will pass directly over the pickup areas of the head. The required degree of height accuracy increases as the tracks become narrower. Table 28-1 shows signal loss due to height errors for several popular tape formats.

Table 28-1. Loss Due to Height Error

Loss		Height Errors in Mils for Various Track Widths*				
dB loss	**% loss**	**84 mil**	**70 mil**	**43 mil**	**37 mil**	**21 mil**
0.1	1.14	0.96	0.80	0.49	0.42	0.24
0.3	3.39	2.85	2.38	1.46	1.26	0.71
0.5	5.59	4.70	3.92	2.41	2.07	1.17
1.0	10.87	9.13	7.61	4.68	4.02	2.28

*84 mil—some 2 track ¼ inch stereo
70 mil—4 track ½ inch, 8 track 1 inch, 16 track 2 inch
37 mil—4 track ¼ inch, 24 track 2 inch
43 mil—24 track 2 inch (on some systems)
21 mil—stereo compact cassette, 8 track ¼ inch

For a tape guide to position the tape accurately, the tape must fit snugly into the guide, but the guide must not squeeze the tape edges. The typical manufacturing tolerances of 2 mils to 4 mils (50 µm to 100 µm) on tape width and 1–3 mils (25–75 µm) on tape guide width result in a loose fit for many rolls of tape.

Sources of height error also include fixed errors in head and guide height and core placement tolerances within the head. A good alignment should contain no more than 1 mil (25 µm) combined error for the head and guides, but this degree of accuracy requires the use of optical measurement devices that are not commonly available in a recording studio. Typical maintenance shop practices will yield errors in the range of 2–3 mils (50–75 µm). When this alignment error is added to a typical core placement error of 1 mil (25 µm) and a tape guide clearance error of 2 mils (50 µm), the signal loss or variation can easily exceed 1 dB on a 24 track recorder.

A relatively simple method of reducing the sensitivity to height errors is to use different widths for the record and reproduce head core widths. Using either a wide playback head on a narrow recording or a narrow playback head on a wide recording will reduce or eliminate the losses due to height variation. Differing track widths, however, give rise to a common operator error. Setting the normal and sync reproduce levels from a full-track alignment tape, which has signal recorded across the entire width of the tape, will produce a level error on the wider of the two heads. The amount of error, which depends on the ratio of the core widths of the two heads, must be subtracted from the actual meter reading of the wider core to determine the true flux level. For example, a recorder with 37 mil (0.93 mm) record cores and 43 mil (1.08 mm) reproduce cores would be set to read 0 VU in sync playback and +1.3 VU in normal playback from a full-track alignment tape.

28.2.4.2 Head Azimuth

Not only must the tape passing across the head be at the correct height, but also the recorded signal on the tape must be parallel to the pickup gap in the reproduce head. Any angular error is referred to as *azimuth error*. Table 28-2 gives the amount of signal loss due to azimuth error for a 15 kHz signal at 15 in/s (38 cm/s), a 1 mil (25 µm) wavelength λ.

For a typical professional recorder with guides spaced 6 inches (15.2 cm) apart, the worst case combination of guide and tape sizes could produce a maximum dynamic guiding error of ±5 mils (125 µm) at each guide, yielding an azimuth error of ±0.1 or ±6 min. This error would generate a signal fluctuation of 3 dB for a 250 mil (6.35 mm) track width as indicated in Table 28-3. Overlapping heads or tracks offer no azimuth loss improvement.

Table 28-2. Azimuth Loss at 15 kHz, 15 in/s (1 mil wavelength)

	Azimuth Error (minutes)			
dB loss	**250 mil**	**70 mil**	**37 mil**	**21 mil**
0.5	2.55	9.1	17.3	30.4
1.0	3.60	12.9	24.3	42.8
3.0	6.10	21.7	41.1	72.4
6.0	8.30	27.3	56.0	98.6

$$loss = 20\log\frac{\dfrac{\sin(\pi W \times \tan\alpha)}{\lambda}}{\dfrac{\pi W \times \tan\alpha}{\lambda}}$$

where,
W is the track width,
α is the azimuth error,
λ is the wavelength.

After many years of decreasing track widths to fit more tracks on a tape, there is a resurgence of wider track formats such as 2 tracks on ½ inch or 1 inch tape and 8 tracks on 2 inch tape. Although these formats yield superior SNR and reductions in amplitude modulation noise, the wide tracks can create level stability problems. These wider tracks will have large signal variations at high frequencies if the tape guides permit even small dynamic guiding errors. As a result, when a transport is fitted with wide-track heads, the guiding is also usually modified to hold the tape more accurately.

For multitrack recorders, the time and phase relationship between audio channels that are recorded on separate tracks may be more critical than the level of short-wavelength signals. Azimuth errors contribute to differential timing errors between tracks, since the azimuth tilting causes one track to be reproduced slightly later than the other. As the distance between tracks becomes large, such as for 1 inch and 2 inch (2.5 cm and 5 cm) formats, the timing error becomes critical. A typical method to measure this timing error is to record the same high-frequency signal on two tracks, and then measure the phase difference between tracks. Table 28-4 shows the amount of worst-case phase difference and timing difference at a 1 mil (25 µm) wavelength introduced by a 0.5 dB head azimuth error for the outer pair of tracks.

The magnitude of both the height loss and the azimuth loss could be greatly reduced if the widths of the tape guides and tape matched perfectly. One method to achieve this objective is to use adapting guides with spring-loaded movable flanges so that the guide adjusts itself to the tape width. Some digital audio recorders with numerous very narrow tracks utilize spring-loaded guides to maintain close repeatability of the tape path.

Table 28-3. Errors Due to 0.5 dB Azimuth Error (1 mil wavelength)

Format	Phase Error	Timing Error
¼ inch stereo	151°	0.28 ms
1 inch 8 track	867° (2.4 rotations)	0.16 ms
2 inch 24 track	3500° (9.7 rotations)	0.65 ms

A similar effect can be achieved with fixed- flange guides if a curvature is deliberately introduced into the tape path. Fig. 28-12 illustrates two possible methods to achieve this curvature. Typically, an offset of less than 5 mils (125 µm) is adequate to overcome the worst-case combination of clearance between the tape and guides and the maximum amount of natural bowing in the tape due to slitting and subsequent handling distortions.

Although the dynamic guiding variations are greatly reduced by forcing the tape to maintain a distorted tape path, the increased force applied to the edges of the tape produces new problems. Not only do both the guides and the edges of the tape experience higher wear rates, but scrape flutter is also increased dramatically. The edges of a tape are very rough due to the shearing action used in the tape-slitting process. When these rough edges slide firmly against the distorting guide flange, tape vibrations are excited, producing scrape flutter.

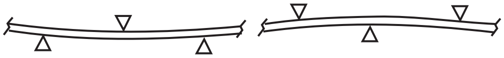
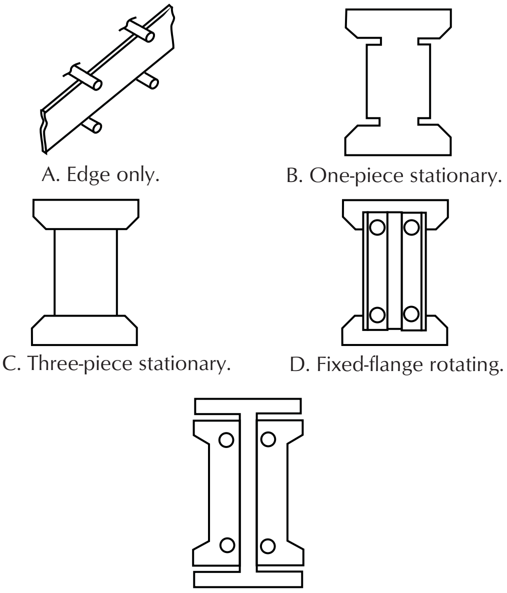


Figure 28-12. Deliberate tape curvature to reduce guiding errors.

28.2.4.3 Tape Guides

Tape guides come in many shapes, sizes, and basic types, as shown in Fig. 28-13. Each guide contains flanges that press against the edges of the tape to steer it. In all cases except the edge-only guide, the tape wraps around the guide to generate stiffness so that the steering force exerted by the flange can move the entire width of the tape and not just buckle the edge. Typically, at least 10° of wrap is required for adequate stiffness.



A. Edge only.

B. One-piece stationary.

C. Three-piece stationary.

D. Fixed-flange rotating.

E. Rotating-flange rotating.

Figure 28-13. Five styles of tape guides.

Rotating guides are generally less effective than stationary guides. Since the tape is in firm contact with the spinning surface of a rotating guide, rather than in sliding contact as with the stationary guide, the force required to slide the tape up or down is determined by the tape tension and the coefficient of *static friction*. The tension component is identical for the stationary guide, but in this case the coefficient of *sliding friction*, which is typically half the static value, is used.

Although both stationary and rotating guides are commonly used in tape transports, rotating guides are slightly more prone to damage the edge of the tape. Guides with large rotating flanges can produce ruffles on the edge of the tape if the tape edge contacts the outer radius of the moving flange. Most guide designs taper the flange to minimize this hazard, but a small flat area at the bottom of the taper is still required if the guide is used for precise tape positioning.

The edge-only guide is very limited in effectiveness since any appreciable force on the edge of the tape may cause the tape to twist rather than move up or down.

28.3 Magnetic Heads

Although magnetic tape is covered in a later section, the following discussion of magnetic heads requires a few very simple assumptions regarding the composition and dimensions of the magnetic tape. First, assume that the magnetic coating consists of microscopic particles of magnetic materials that have been bonded to one surface of a thin plastic backing or substrate. Second, each magnetic particle is assumed to function as a small independent magnet, allowing patterns of varying magnetic polarity and intensity to be stored along the tape. Last, the thickness of the magnetic coating for the audio tape example will be assumed to be 0.6 mil (15 µm).

28.3.1 Geometric Characteristics

Most of the characteristics of magnetic heads are controlled by the geometry of the head and the magnetic tape. Since wavelength on tape is determined by the recorded frequency in hertz and the relative tape-to-head speed, there can be many combinations of frequency and speed that will result in the same effects in a head. For example, the wavelength of a 15 kHz tone on a mastering recorder at 15 in/s will have the same wavelength as a 240 kHz signal on a high-speed tape duplicator running at 240 in/s. The geometric considerations for both applications are identical, despite the 16:1 difference in tape speed.

Not all of the characteristics are geometric, however. Eddy current losses, for example, depend on the frequency in Hz rather than the wavelength.

28.3.1.1 Gap Length Loss

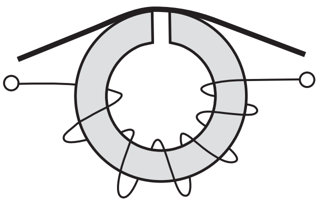
Each of the tiny magnetic particles on the surface of the tape produces a magnetic force or flux in the space surrounding the particle. This invisible magnetic effect, called a *magnetic field*, will interact with other nearby magnetic particles. To measure the strength of this field, a flux concentrator in the form of a reproduce head is scanned along the tape. The resulting electrical output from the head is dependent on the flux pattern recorded on the tape.

The reproduce head must be able to collect flux selectively from a very small span of tape. For example, flux patterns on a compact cassette may be as small as 100 millionths of an inch (100×10^{-6} in or 2.5 µm) in wavelength. To achieve this fine resolution, a small gap must be created in a ring of magnetic material, as shown in Fig. 28-14A.

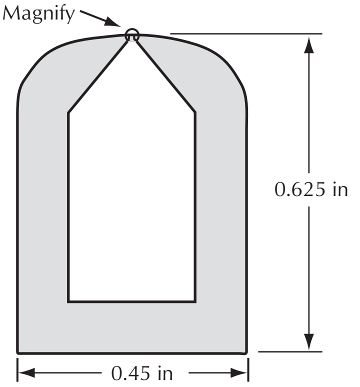
The length of the gap ranges from two ten-thousandths of an inch (2×10^{-4} inch or 5 µm) for studio mastering recorders down to less than 30 millionths of an inch (30×10^{-6} inch or 0.75 µm)—the wavelength of red light for cassette and high-density digital recorders. Since no slicing technique is available to cut accurate gaps that short, the core is usually fabricated as two pole pieces that are fastened together with a shim spacer of the desired dimension inserted in the gap. Fig. 28-14B shows a typical studio head core drawn full size, with the critical gap area at the pole tips and adjacent tape magnified in Fig. 28-14C.

The operation of the gap, which serves as a sensing aperture, can be analyzed in terms of a flux pickup focused at the surface of the tape. The amount of flux picked up by the core, and thus made available to generate an output voltage in the winding, is determined by the *net* magnetic flux from pole tip to pole tip across the gap area. If the tape segment at the gap consists of a strong magnetization of only one polarity, the flux in the core will be maximized. If, on the other hand, the segment contains two strong portions of opposite polarity that cancel each other, the net flux in the core will be zero.

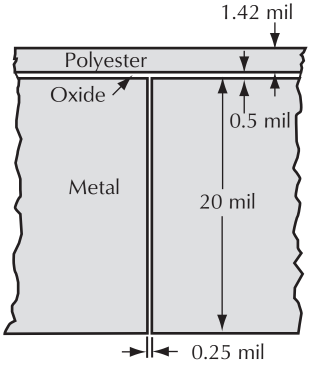
The efficiency of the gap due to this averaging effect is illustrated in Fig. 28-15. The output of the head declines, slowly at first, and then quite rapidly to zero as the wavelength decreases to the length of the gap. As the gap length becomes longer than the wavelength, an output of opposite polarity appears. When the wavelength drops to half the gap length, another null will occur. This pattern of diminishing peaks of alternating polarity is repeated over and over, with nulls occurring



A. A small gap in a ring of magnetic material.

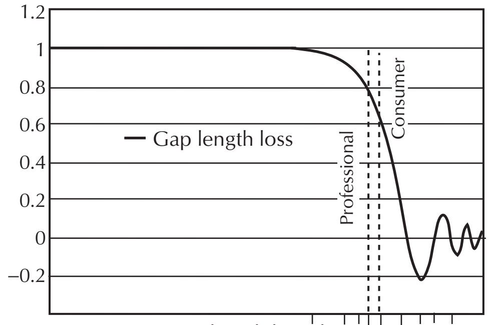


B. Typical studio head.

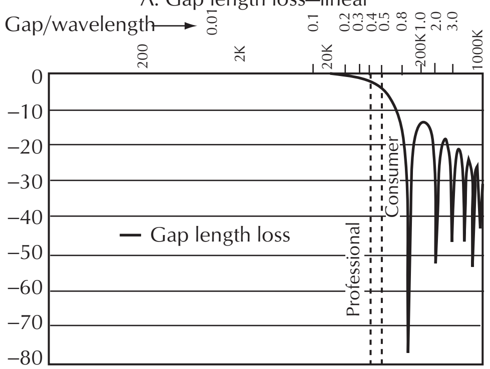


C. Magnified critical gap and adjacent tape.

Figure 28-14. Ideal and practical magnetic heads.



A. Gap length loss—linear



B. Gap length loss—dB

Figure 28-15. Loss due to gap length and the ratio of gap length to wavelength.

at each wavelength that produces an odd or even number of complete cycles in the gap.

The example recorder at 15 in/s has only 1 dB of gap length loss at 40 kHz and 3 dB at 67 kHz, certainly not a dominating loss. At 30 in/s these losses become even more insignificant, with the –1 dB and –3 dB frequencies doubling to 80 kHz and 134 kHz.

Audio recorders are seldom designed to operate beyond the dashed lines shown in Fig. 28-15. With this constraint, gap length loss for professional machines can be held below 1 dB or 2 dB by choosing an appropriate gap length for a given application and minimum wavelength. Mastering recorders operating at 15 in/s and 30 in/s and broadcast machines operating at 7.5 in/s have playback gaps ranging from 100–200 µinch, compact cassette machines operating at 1⅞ in/s have gaps of 30–60 µinch.

Mastering recorders may also use the record head for playback in the *sync* mode. Since the record heads may have gaps ranging from 250 µinch to 1000 µinch, the sync response may suffer significant high-end loss. For example, a 1950s vintage recorder with a 1000 µinch record gap will reach its first null at 15 kHz for a tape speed of 15 in/s. As sync response became more important in the mid 1960s the recorder manufacturers tightened up the record gaps to 350 µinch or less to improve sync response.

If the gap length is inferred from the first measured null, this *effective gap length* may be 10% to 15% longer than the mechanical gap determined by the shim. Various proposed explanations include magnetic degradation of the inner surfaces of the pole tips due to manufacturing stresses and pole tip saturation. When in doubt, add 10% to the optically measured length or shift the response points down to 91% (1/1.1) of the theoretical values. For the ATR100 example, the −1 dB and –3 dB points would shift to 36 kHz and 61 kHz.

Use of an excessively short gap will cause an additional loss in overall head sensitivity due to shunted flux that jumps the gap rather than traveling through the core, as shown in Fig. 28-16. For this reason, the reproduce head gap length is usually chosen to give the largest acceptable loss at the shortest expected wavelength.

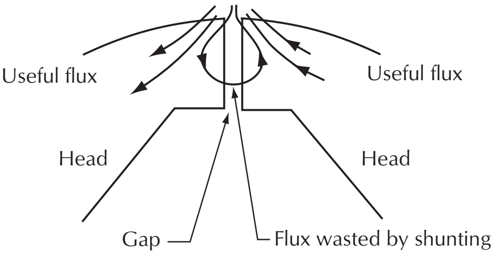


Figure 28-16. Gap shunting loss.

28.3.2 Spacing Losses and Thickness

The recording process magnetically aligns groups of the tiny randomly oriented magnetic particles so that they act as if they were a single larger particle. We could visualize these groups as little bar magnets that have dimensions determined by the tape and signal. The track width defines the vertical direction and the tape coating thickness sets the depth. The length is determined by the wavelength of the recorded signal. To simplify the example, assume that a 1.5 kHz square wave is recorded at a tape speed of 15 in/s, yielding a wavelength of 10 mils or 0.010 in. The recorded image is similar to a series of bar magnets each 5 mils long with alternating polarity.

Actually, gap length loss and shunting loss are only a part of what determines the performance of an audio recorder. The most critical parameter is the relative thickness of the magnetic coating on the tape. The ratio of tape thickness to the shortest wavelength to be recorded has a profound effect on the frequency response, maximum output, noise, and signal-level fluctuations.

The magnetic particles at the surface of the tape are very tightly coupled to the core of the head, producing a maximum amount of playback flux in the core. Particles that are buried below the surface of the tape, however, produce a weaker flux in the core. The amount of flux that is lost depends on the spacing distance and the wavelength—just as a small font size is more difficult to read at a distance than a larger font. An approximate expression for this spacing loss is

$$Spacing\ loss_{dB} = 55 \times \frac{distance}{wavelength}. \tag{28-4}$$

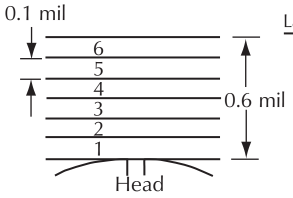
One example of the use of this spacing loss formula is to determine the playback signal loss due to a piece of dirt on the surface of a reproduce head. Assuming a typical recording studio tape speed of 15 in/s (38 cm/s), a dirt speck only 0.0001 in (2.5 µm) high will produce losses at the following frequencies of

$$\begin{aligned} \text{150 Hz spacing loss} &= 55 \times \frac{0.0001}{\frac{15}{150}} \\ &= 0.055\text{ dB} \\ \text{1500 Hz spacing loss} &= 0.55\text{ dB} \\ \text{15 kHz spacing loss} &= 5.5\text{ dB} \end{aligned}$$

Note that this seemingly insignificant dirt particle has produced a serious loss in high frequencies.

Spacing loss due to dirt is not the major problem created by the "nearsightedness" of the gap since proper head cleaning will keep spacing distances to less than 10^{-5} inch, which is (0.25 µm), producing virtually no error at studio tape speeds. The problem is eight times more severe for cassette speeds of 1⅞ in/s (4.8 mm/s).

The major spacing problem arises within the tape itself since the magnetic coating thickness spaces most of the particles away from the head with other particles. Consider the tape to be composed of several independent layers of oxide, as shown in Fig. 28-17. The average spacing loss for each layer, calculated using the midpoint of each layer to determine the spacing distance, is tabulated for the example with a typical 0.6 mil (15 µm) coating thickness.



Layer	Spacing	15 kHz loss
1	0.05	–2.75 dB
2	0.15	–8.25 dB
3	0.25	–13.75 dB
4	0.35	–19.25 dB
5	0.45	–24.75 dB
6	0.55	–30.25 dB
	Total output	+ 3.63 dB

Figure 28-17. Tape thickness loss.

The contributions of layers 2 through 6 fall off so rapidly due to spacing loss that their *combined* contribution is only equal to layer 1 by itself at this wavelength. Indeed, shaving off layer 6, which constitutes 17% of the coating thickness, would produce a loss of only 2% or 0.18 dB in output at this wavelength.

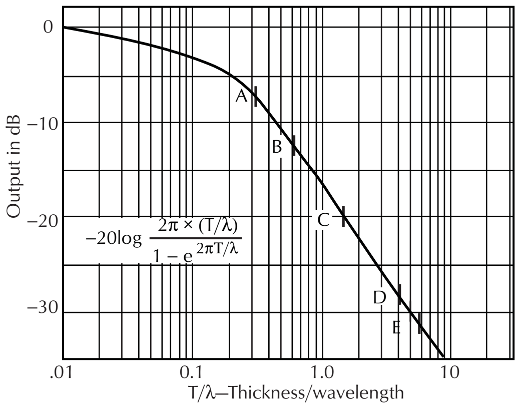
This coating thickness loss can be expressed as

$$Coating\ thickness\ loss_{dB} = 20\log\frac{x}{1-e^{-x}} \tag{28-5}$$

where,

x is 2 π × thickness/wavelength.

Although this expression yields a drop of 6 dB per octave, as shown in Fig. 28-18, this curve is not the same shape as the response of a low-pass filter made from a resistor and capacitor. The response in Fig. 28-18 is down 4 dB at the intersection of the asymptotes rather than the typical 3 dB for a single pole RC filter. This difference in shapes means that a simple RC boost circuit will not properly correct for the thickness loss. Depending on the choice of RC boost frequencies, the difference in shape will produce an error of 0.5–1.0 dB in the midband response.



A. 15 kHz 30 in/s T = 0.65 mil (−7.4 dB)
B. 15 kHz 15 in/s T = 0.65 mil (−12.4 dB)
C. 15 kHz $1^7/_8$ in/s T = 0.2 mil (−20.0 dB)
D. 15 kHz $1^7/_8$ in/s T = 0.65 mil (−30.3 dB)
E. Digital 30 kbit/in T = 0.2 mil (−31.5 dB)

Figure 28-18. Loss due to ratio of coating thickness to wavelength.

28.3.2.1 Equalization Boosts

This thickness loss of Fig. 28-22, must be corrected by applying compensating boosts in either the record or reproduce circuitry. Although this loss is a playback deficiency, the choice of whether to correct the loss during record or playback is somewhat arbitrary. The amount of record boost is limited by the magnetic saturation characteristics of the tape; playback boost is limited by the high-frequency noise characteristics of the tape and the reproduce head and associated circuitry.

The minimum amount of boosting required to achieve flat response can be considered to be a necessary equalization. The industry has developed a set of internationally recognized standards to promote compatibility of tapes. Each standard deals with the necessary and discretionary equalizations to define the exact characteristics of the recorded tape. Using the tape flux characteristics as a standard implicitly specifies the partitioning of equalizations between the recording and reproducing functions. Table 28-4 lists the commonly encountered standards.

Unlike the absolute nature of the reproduce characteristics, the record characteristics of the recorder must have enough flexibility to accommodate a number of different tape sensitivities and frequency characteristics. Once the reproduce section has been calibrated to the standard with a standard alignment tape, all further adjustments are to produce a recorded tape on the machine that accurately matches the standard tape.

The amount of thickness loss can always be reduced by utilizing thinner coatings, but any decrease in thickness also causes an equal drop in low- and midfrequency output and SNR. To preserve the existing standards, the tendency has been to adjust the coating thickness of new tapes to emulate the high-frequency losses of the older tape types while trying to achieve maximum low- frequency output. This somewhat self-defeating strategy has been overcome in recent thin-coat high-energy tapes

Table 28-4. Common Tape Record-Playback Equipment Equalization Standards.

		Transition Frequencies and Time Constants				
Standard	**Type Tape**	**1⅞** ips	**3¾** ips	**7½** ips	**15** ips	**30** ips
IEC	Fe_2O_3	100/1326 Hz	50/1768 Hz	0/2274 Hz	0/4547 Hz	
		1590/120 µs	3180/90 µs	∞/70 µs	∞/35 µs	
	Metal	100/2274 Hz				
		1590/70 µs				
NAB		50/1768 Hz	50/1768 Hz	50/3180 Hz	50/3180 Hz	
		3180/90 µs	3180/90 µs	3180/50 µs	3180/50 µs	
AES						0/9095 Hz ∞/17.5 µs

that retain the low-frequency output capability of older tapes, but utilize new equalization curves optimized for the new tape thickness.

28.3.2.2 Fringing

Spacing loss is evident not only at high frequencies, but it also shows up at very long wavelengths. Magnetic information that is recorded off to the sides or fringes of the area normally scanned by the reproduce head core will begin to be sensed if the wavelength becomes longer than the separation distance. At studio operating speeds of 15 in/s and 30 in/s (38 cm/s and 76 cm/s), for which low-frequency wavelengths reach ½ inch and 1 inch (12.7 mm and 25 mm), this fringing leakage becomes very evident. For example, for the case of an oversized record track mentioned in Section 28.2.4.1, the signal level may rise by the ratio of the track width to the core width. A typical 24 track format on 2 inch (51 mm) tape would encounter a 1.3 dB rise at frequencies below 500 Hz for record and reproduce cores of 43 mils and 37 mils (1.1 mm and 0.9 mm).

A similar case arises when alignment tapes made with a single full-width record head are utilized for level and response checks. The sideways fringing will produce significant level and response errors. The actual amounts of error depend on both the track format and the playback head design. Some alignment tape manufacturers roll off the low frequencies in an attempt to offset the rise in a nominal head, but the amount of this fringing compensation is not absolutely correct for all head designs.

One additional pitfall to be avoided is the fringing differences between center tracks and edge tracks. Since the edge cores run very near the physical edge of the tape, these cores sense only one-half the amount of fringing flux sensed by the inner cores. During a frequency check from a full-width alignment tape, the two edge tracks should therefore be slightly lower in output at the low frequencies than the remainder of the tracks.

28.3.2.3 Contour Effect

At very low frequencies, the wavelength of the recorded signal may become as long as the magnetic core of the playback head. These long wavelengths enter the core at the gap and at the sides and rear of the core. The resulting flux in the core will consist of the desired flux from the gap plus additions and/or subtractions of the fringing flux leaking into the core at the sides and back. The voltage output of the head, which is dependent on the net flux coupled into the windings, will undulate at low frequencies as the wavelengths create varying levels of constructive and destructive interference due to the fringing flux.

The response curve in Fig. 28-19 illustrates the nature of the undulations or *head bumps* for a typical mastering recorder at 15 in/s (38 cm/s) and 30 in/s (76 cm/s) using a reproduce head that has a 0.5 inch (12 mm) core face. Two well-defined head bumps are usually evident for such mastering heads. The bumps shift up an octave in frequency for each doubling of tape speed, creating an even more severe problem at 30 in/s (76 cm/s).

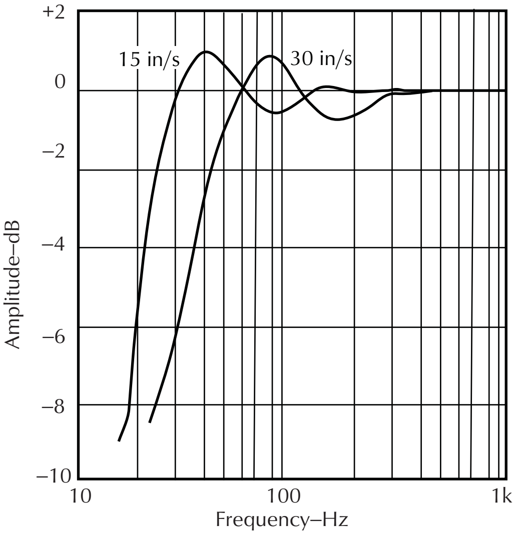


Figure 28-19. Contour effect. Courtesy Sony Corporation of America.

Heads with either very small cores or only a small window in the head shielding at the gap area can produce numerous ripples in the low-frequency response. Such heads should be avoided unless the tape speed is slow enough to avoid serious problems within the normal band of audio frequencies.

The exact shape of the head bumps is determined by the size and shape of the reproduce core, surrounding shielding material, and angle of wrap of the tape. Since the user cannot adjust these parameters during the normal alignment procedure, the bumps can only be modified by adding an outboard equalizer, which cancels the bumps with an inverse response curve.

Recent improvements in the control of head bumps has reduced the magnitude of the bumps in present-day mastering recorders to less than 1 dB peak-to-peak at 15 in/s (38 cm/s) and 1.5 dB peak-to-peak at 30 in/s

(76 cm/s). Beware that this level of error will be introduced each time the tape is rerecorded during mixdown and subsequent protection copying. The total error can easily reach 5 dB or more for a typical sequence of operations.

28.3.2.4 Crosstalk

Fringing also produces playback signal leakage or crosstalk between adjacent tracks at long wavelengths. The unused area or guard bands between the cores of the head, which are nearly equal in width to the recorded track, usually provide enough of a physical gap to prevent flux from spilling from one track to the next. At long wavelengths, however, the fringing flux will jump the guard band, producing low-frequency crosstalk.

The crosstalk component due to fringing will initially decrease as the frequency is increased, but at midband the decrease will eventually bottom out. The remaining residual level of crosstalk is not due to fringing, but it is a direct transformer-like coupling of leakage flux between the adjacent cores in either the record or reproduce head. A layer of magnetic shielding material is typically placed between the cores of the head as a crosstalk shield to reduce this flux leakage.

28.3.3 Frequency Characteristics

28.3.3.1 Inductive Rise

Up to this point, most of the losses and response anomalies have been governed by the wavelength performance of the interface between the tape and the head. An additional set of characteristics due to the internal frequency-dependent operation of the head must also be considered.

The most striking characteristic in the frequency response of a conventional coil-and-core playback head is a continuous 6 dB/octave rise in output voltage with rising frequency. The core and winding of the head form an inductor in which the output voltage is proportional to the rate of change of the flux in the core as seen in the equation

$$\text{head } V_{out} = N\frac{\Delta\phi}{\Delta t} \tag{28-6}$$

where,

N is the number of turns in the winding,

$\Delta\phi$ is the change in flux,

Δt is the time interval.

Any ratio of the form $\Delta x/\Delta t$ is called a differential with respect to time, and the device creating this rate of change is called a *differentiator*.

If a sine-wave signal of frequency f is used for testing the output voltage of a head, the voltage expression can be further simplified to

$$\text{head } V_{out} = 2\pi N f \times flux_{max} \tag{28-7}$$

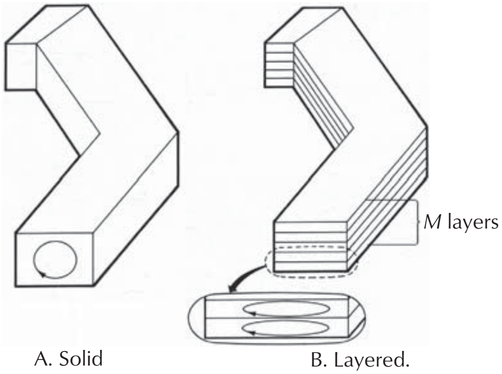
28.3.3.2 Hysteresis Loss

The constantly changing magnetic flux in the core of the reproduce head gives rise to losses within the core of the head. One source of these losses is the amount of energy that is required to change the magnetization state of the core material. Every time the flux in the core reverses polarity, a small amount of energy is lost in overcoming the magnetic memory or hysteresis of the core material. The hysteresis power loss increases with both increasing flux magnitude and frequency.

28.3.3.3 Eddy Current Loss

The changing core flux generates a voltage not only in the winding of the head, but also within the core itself. If the core is metallic, this voltage will cause a current to flow within the core, as shown in Fig. 28-20. The core currents, referred to as *eddy currents* because of their similarity to swirling eddies in a stream of water, dissipate energy that should be going to the reproduce signal.

The amount of power (P) dissipated in the eddy currents is given by the general power equation:



A. Solid B. Layered.

Figure 28-20. Eddy current.

$$P = \frac{V^2}{R} = I^2R \tag{28-8}$$

The previous discussion on the inductive rise of voltage with increasing frequency in the reproduce head also applies to these eddy components, producing a rapid rise in eddy current power loss.

The eddy currents of the solid core of Fig. 28-20A rise to an unacceptable level even before the upper limits of the audio band are reached. Fortunately, this drastic loss can be decreased by dividing the core into many thin insulated layers or laminations, as shown in Fig. 28-20B. For *M* laminations, each lamination would generate only 1/*M* of the core voltage and 1/*M*^{2} of the loss power produced by a solid core. The core resistance for each lamination drops only slightly since the width of the lamination remains unchanged. The net improvement for *M* laminations is a 1/*M* reduction in the eddy current power loss. (Professional audio heads, which are typically constructed with laminations 2 mils (50 µm) thick, will contain 20 to 120 laminations per track, depending on the track width.) Reducing the core size and using high-resistivity core materials such as ferrites can achieve even further improvements.

28.3.4 Combined Characteristics

Fig. 28-21 illustrates some of the individual and composite effects of the foregoing reproduce head characteristics. The constant 6 dB/octave inductive rise of the head has been omitted in the illustration to accentuate the undesired departures from flat response.

Curves A, B, and C illustrate the gap length, tape thickness, and spacing losses, respectively.

Curve D represents a typical resonant rise due to head inductance and the capacitance of the head cable and head winding. The playback amplifier high-frequency response boost dictated by the National Association of Broadcasters (NAB) equalization standard for 15 in/s (38 cm/s) is represented by curve E. The combination of all of these effects in curve F yields a response that is flat within ±1 dB. This simplified model does not include relatively minor contributions at mastering speeds due to eddy currents and hysteresis, self-demagnetization effects, recording equalization, and the effects of nonuniform distribution of recorded flux due to coating thickness. In spite of these omissions, the dominant nature of the coating thickness loss is readily apparent. The equalization standards have been chosen primarily to offset this thickness loss.

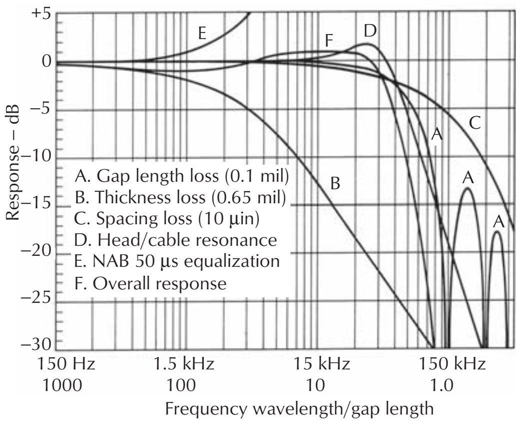


Figure 28-21. Playback loss components at 15 in/s for 0.65 mil coating thickness.

The composite curve F represents the overall playback performance from an ideal tape of finite thickness. All the indicated response anomalies within the audio band must be either corrected or tolerated. In some cases, one effect can be used to offset others, such as shaping the resonance curve to compensate for the gap length loss. (Unlike the resonance, the gap length loss increases with decreasing tape speed, upsetting the compensation at lower tape speeds.)

28.3.5 Noise

The useful range of signal levels that pass through the tape recorder is limited by the maximum signal at which all the magnetic tape particles become completely magnetized or saturated and also by the amount of noise that remains when the input signal is removed. Noise in tape recorders has many sources; the electronics, the tape, and the heads themselves all contribute to the residual noise.

The distortion content of the signal from a tape recorder rises so dramatically near tape saturation that the normal operating range must be limited to less-than-maximum levels. For the purpose of specifying and comparing tape recorders, the distortion-free maximum operating level is typically considered to be the output signal level at which the *THD*, which is dominated by third harmonic and other odd components, reaches 3%. The ratio of the level for 3% THD at medium wavelength to the residual noise is defined as the SNR of the recorder.

28.3.5.1 Track Width

The second factor is the loss in SNR in narrow-track consumer tape formats due to the dissimilar ways that random noise sources and coherent signals increase. The noise due to the tape, heads, and electronics is a random combination of many small independent noise bursts. If two equal and independent random noise sources of this type are added together, the noise *power* is doubled, producing an increase of 3 dB on a voltmeter.

Coherent sources, on the other hand, are merely duplicates of the same waveform. If two identical sources are added together, the value at each point on the output waveform is exactly twice the value of either of the input waveforms. In this case the output *voltage* is doubled, or a 6 dB increase.

Consider the case of two tracks of a tape recorder that have recorded the same signal. If the output signals of the two tracks are added, the noise will add randomly and the signals will add coherently. The combined tracks have 6 dB more signal and 3 dB more noise, yielding a net SNR improvement of 3 dB. Using a single track of double the original track width would produce the same result if the noise sources were statistically independent in nature.

The tape noise will follow the 3 dB per doubling rate if the reproduce amplifier noise is less than the tape noise. The reproduce amplifier noise typically remains nearly constant regardless of track width of the head. The apparent noise will vary, however, as the gain of the amplifier is adjusted to compensate for changes in the head output due to increased or decreased track width. When tracks are made narrower, the amplifier noise that functions as a coherent source will eventually dominate the tape noise, creating a signal-to-noise loss of 6 dB per halving of tape width.

Fig. 28-26 compares the output voltage and signal-to-noise variation for various track widths, assuming that all noise sources are truly random for a noiseless preamplifier and a typical preamplifier. When the amplifier noise begins to dominate the other noise sources, there is a rapid loss of SNR with decreasing track width.

28.3.5.2 Thermal Noise

Both the core and winding of the reproduce head contribute random noise to the output signal. For the winding, the noise source is due to the thermal agitation of the atoms in the copper wire. The amount of thermal noise is given by the expression

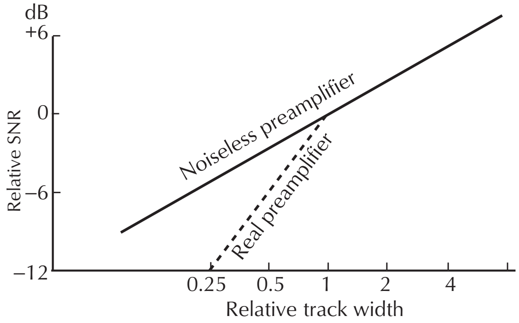


Figure 28-22. Noise changes with track width.

$$\begin{aligned} TN &= \sqrt{4KTRB} \\ &= 1.82 \times 10^{-8}\sqrt{R} \text{ volts for a 20 kHz} \\ &\quad \text{bandwidth at room temperature} \end{aligned} \tag{28-9}$$

where,

TN is the thermal noise,

K is the Boltzmann's constant (1.38×10^{-23} joules/K),

T is the absolute temperature in kelvin,

R is the resistance in ohms,

B is the measurement bandwidth in hertz.

A 100 Ω resistor will produce 0.182 μV of noise voltage. Depending on the core size and number of turns, a playback head may exhibit a resistance from 10–1000 Ω, yielding thermal noise contributions of 0.06–0.6 μV. The increase in noise due to more turns of finer wire in high-inductance heads is offset by a rise in head output voltage, producing little net change in SNR.

28.3.5.3 Barkhausen Noise

Another major noise source is *Barkhausen noise*, a noise due to jumps in the magnetic boundaries of the core material. The core metal consists of a collection of many microscopic magnetic zones or domains. When a magnetic field is applied to the core, the boundaries or walls of the domains will change as small domains merge to form larger domains. This merging occurs in discrete steps since the small domains act as single units that must each merge completely in one jump. The resulting step change in the magnetic field generates a noise burst in the head winding. Since the core contains millions of constantly switching domains, a statistically independent random noise is generated. Reducing the size of the basic domains will decrease the amplitude of the Barkhausen noise.

28.3.5.4 Magnetostrictive Noise

The magnetic core material also exhibits magnetostriction—a change in magnetic field due to stress. The microscopically rough surface of the magnetic tape will therefore produce a small magnetic field change in the core as the tape slides across the head. This field change generates a magnetostrictive noise component in the winding.

Both the Barkhausen and magnetostrictive noises are absent when no tape is moving over the surface of the head. The residual standby noise, which is measured under these conditions, is the absolute noise floor for the reproduce head and amplifier. The comparison of this standby noise level with the bulk-erased and biased noise levels is covered in the test and maintenance section.

28.3.6 Record Heads

The magnetic core and gap of a reproduce head obey the principle of reciprocity, which states that the roles of an excitation source and sensor can be interchanged. For a head used in the reproduce mode, external flux at the gap produces a voltage across the head winding. If, instead, a voltage is applied to the head winding, a concentrated external flux field will be generated at the gap and can be used to record a signal on a piece of moving tape.

The shape and strength of the magnetic field at the gap is the basis for the operation of a recording head. The flux generated in the core by the current in the winding must jump across the gap to complete a closed magnetic path. The gap, which is a very poor magnetic path compared to the core, produces an obstruction that forces the flux to spread sideways, as shown in Fig. 28-23.

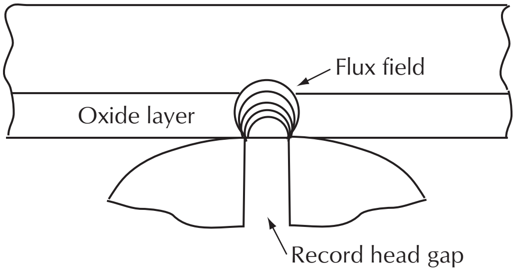
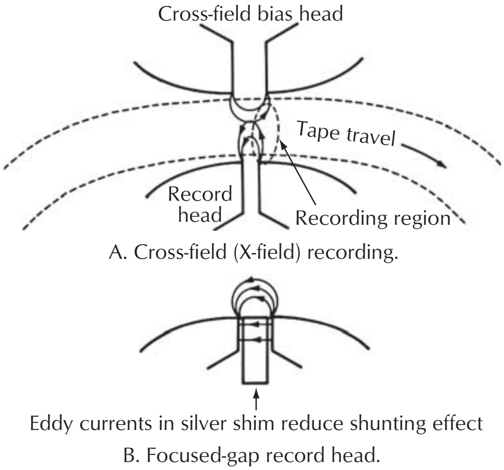


Figure 28-23. Record head flux field.

An analogous situation occurs with a crowd of people moving down a hallway. If the hallway widens for a crossing corridor or small lobby, the crowd will broaden out into the open area and then narrow down again to reenter the continuation of the hallway. The broadening will increase if the pressure within the hallway should increase due to an emergency such as a fire. The stress is greatest at the transitions between the wide and narrow spaces since this is where people are squeezing to try to change the shape of the flow.

A magnetic tape passing over a record head gap experiences a similar buildup and decline in the magnetic recording field as it moves across the gap. To produce a permanent recording on the tape, the flux must first rise to a level sufficient to overcome the magnetic memory force of the tape, which normally keeps the magnetic particles on the tape from changing state spontaneously. In the central zone of complete excitation, the tape particles will follow any change in the input signal driving the head. As the tape particles exit the strong central zone, a well-defined point will be reached at which the driving flux drops below the memory force, leaving a fixed magnetic image impressed on the tape. This transition region in which the image freezes at the trailing edge of the gap is called the *trapping plane.*

The shape of the trapping plane depends primarily on the gap size and the thickness and magnetic characteristics of the tape. Since trapping planes that are narrow and vertical will produce short-wavelength recordings that are more easily reproduced, several techniques have been developed to sharpen the transition zone, as shown in Fig. 28-24.



A. Cross-field (X-field) recording.

B. Focused-gap record head.

Figure 28-24. Focused-gap and cross-field (X-field) head.

The focused-gap technique in Fig. 28-24B uses a highly conductive gap shim made of silver to serve as a barrier to flux jumping straight across the gap. Eddy currents in the shim force the flux away from the shim, squirting the flux deeper into the tape. The reduction in shunted flux raises the efficiency of the head by requiring less drive power.

The conductive shim is only effective at high frequencies at which large eddy currents are generated in the shim. As a result, focused gap recorders utilize bias frequencies that are approximately ten times higher than conventional systems.

In practical use the silver shim proved to be a major problem because the soft silver would smear onto the trailing pole piece of the head and short the head's laminations together.

A second technique, which yields similar results, is the *crossed field* or *X-field*, Fig. 28-24A. This method typically places a second bias-only head on the back side of the tape to create a shaped bias flux field jumping from one head to the other.

28.3.6.1 Biased or Anhysteretic Recording

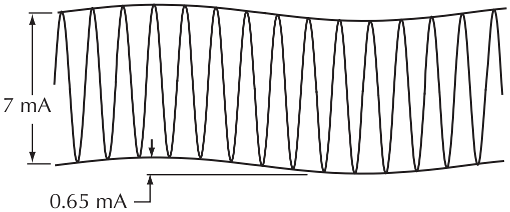
The magnetization of the tape particles is not easily changed due to the memory force or hysteresis of the particles. In fact, the particles have a form of inertia that must be overcome if a linear transfer is to be achieved.

If a rapidly varying signal of sufficient amplitude to just begin magnetizing the particles is added to the audio flux signal, the magnetic particles will more readily conform to changes in the audio waveform. The high-frequency biasing signal produces a hysteresis-free or anhysteretic recording.

Fig. 28-25 shows a typical waveform of the current in a low impedance Ampex record head that is recording 10 kHz at a level of 250 nW/m. The bias component of 7 $\text{mA}_{\text{p-p}}$ is approximately ten times larger than the 10 kHz component at 650 μA. (The voltage waveform across the record head would be totally dominated by the bias component due to the 6 dB/octave rise in head impedance with increasing frequency, in this case 35 V of bias versus 500 mV of 10 kHz or 70:1.)

The audio and bias signals must be added together in a linear manner without generating any of the sidebands that are present in either amplitude or frequency modulation techniques. The short- wavelength bias signal can therefore be easily filtered out during playback by the gap and thickness losses so that only the audio signal remains. (The high level of bias signal transformer crosstalk that is present during sync/overdub operation requires a sharp notch filter in the playback preamplifier to remove the bias signal.)

Typical bias frequencies range from 100 kHz for slow-speed recorders to over 10 MHz for high-speed tape duplicators. Although high bias frequencies are desirable to permit easy filtering and thorough tape excitation, a practical upper limit for mastering recorders is reached at 500 kHz due to a combination of increased eddy current and hysteresis losses in the core and the increase in bias drive voltage required due to the inductance of the head.



Record head current 15 cycles of bias @150 kHz for 1 cycle of audio @10 kHz. Amplitude ratio is 10:1

Figure 28-25. Record head current.

Head losses can be reduced by using a very small core to reduce hysteresis losses and by choosing either thin laminations or a ferrite material to reduce eddy current losses. If, however, the record head will also be used for the reproduce function during sync/overdub, a small core will cause serious long-wavelength contour effects. The compromise *hammerhead* design shown in Fig. 28-26 improves the playback performance of the small core by adding extensions to the face of the core. The tips function only to play back low-frequency signals for which core losses are insignificant.

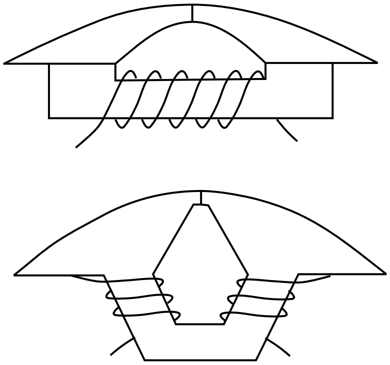


Figure 28-26. Hammerhead cores.

The bias voltage required to drive a record head doubles each time the bias frequency is doubled due to the inductance of the record head. To keep the required bias voltage within the range of common integrated circuits, the inductance can be lowered either by reducing the number of turns in the winding or by lengthening the gap. Reducing the number of turns once again degrades the sync/overdub performance by reducing the playback voltage generated by the head.

Heads with very low inductance typically require a step-up transformer to achieve adequate playback SNRs, but the transformer will also contribute some additional small amounts of distortion, noise, and frequency response anomalies.

Lengthening the record head gap will reduce shunting and give better bias penetration into the tape, but the short-wavelength sync/overdub response will suffer greatly.

A more straightforward approach to optimize the record head for both recording and playback is to use separate flux paths or windings for each of the functions. One simple method of switching windings and flux paths is to use parallel paths that can be selectively blocked. As shown in Fig. 28-27, when the high-inductance playback winding is shorted, flux will be blocked from the playback shunt magnetic leg of the core, effectively eliminating this path and thereby forcing all of the flux from the low-impedance bias winding to the front of the head. During reproduce, when the bias winding is shorted, the flux picked up from the tape will pass only through the reproduce winding. Although the cost of this dual-winding head is significantly higher than for a conventional single-path design, each coil can be optimized for its intended function without the need for compromise, yielding playback-to-record inductance ratios of up to 1000:1.

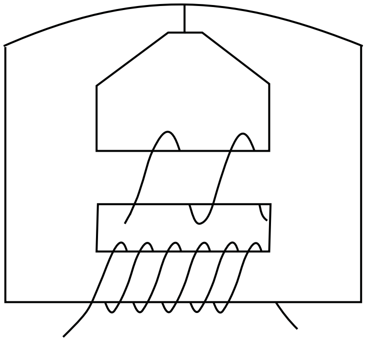


Figure 28-27. Dual winding record head.

28.3.7 Erase Heads

A major advantage of magnetic tape recording is the ability to erase easily and reuse the magnetic tape. Although physical wear may eventually degrade the performance of the tape, the magnetic properties of the tape never wear out.

Erasure of the tape can be accomplished by remagnetizing the tape with either a very strong static field or a very strong alternating field. For audio applications the alternating field, which produces a completely random flux pattern that is very quiet, is used exclusively.

A very large electromagnet, known as a *bulk eraser* or *degausser*, is used for rapid erasure of an entire reel of tape, Fig. 28-28. The coil and core of the degausser are similar to a large recording head. The very strong flux field created across the eraser gap penetrates the magnetic tape, driving the magnetic particles to complete saturation. Any magnetic patterns on the tape are completely erased when all the tape particles are alternately saturated in one direction and then the other by the changing field.

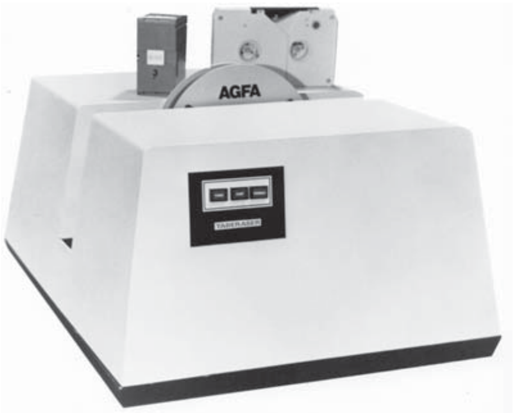


Figure 28-28. Tape eraser. Courtesy Taber Manufacturing & Engineering Co.

To leave the tape in a neutral stage, the strength of the erasing field must gradually decrease from hard saturation to zero. A common technique is to move the tape slowly away from the eraser so that the 60 Hz excitation field will drop gradually from one cycle to the next.

A few degausser models include control devices that gradually reduce the current in the eraser coil to zero, thereby eliminating the need for the operator to move the reel. Other models contain motor-driven actuators that slowly remove the tape from the field automatically.

A dc current or permanent magnet can also be used to erase unwanted signals from the tape, but the tape particles will not be left in a neutral state. A dc-erased tape will usually produce a very noisy recording that contains high levels of even-order harmonic distortion components.

Selective erasure of small portions of a reel of tape requires the use of an erase head on the tape recorder. The function of the head is similar to the bulk eraser in that the tape is slowly withdrawn from a saturating ac flux field. For a tape speed of 30 in/s (76 cm/s) and a typical decay length of the erase head field of 0.005 in (125 µm), the drop from saturation to zero occurs in 0.17 ms. If the tape is to experience at least 20 complete cycles during the decay, the erase frequency must be

$$\begin{aligned} f_{erase} &= \frac{number\ of\ cycles}{decay\ of\ time} \\ &= \frac{20}{0.0017} \qquad (28\text{-}10) \\ &= 120\text{ kHz.} \end{aligned}$$

A conventional coil-and-core head that has a very long gap can produce the long flux field required for erasure. Although such heads will produce approximately 50 dB of erasure, some of the original signal will still remain. A second pass over the erase head will provide the additional erasure that is required to erase the unwanted signal completely.

The reason for the incomplete erasure is a phenomenon known as *gap jumping*. As the tape leaves the saturation zone of the erase gap, the flux level experienced by the tape particles will pass through a level that creates a recording zone similar to the trapping plane of the record head. Any audio variations in the erase field would be recorded at this point. Such audio variations are created by the unerased program that is starting to enter the erase field at the other side of the gap. This incoming flux adds to the erase head flux, creating an unwanted recording at the trailing edge of the erase field just as if the audio signal had jumped across the gap.

Complete erasure can be achieved without multiple passes if the erase head contains two magnetically isolated gaps. The tape is erased by the first gap, and then immediately reerased by the second gap. A wide center spacer isolates the two gaps so that flux cannot jump both gaps.

Although both bulk erasers and erase heads are capable of completely erasing all recorded material from a tape, the residual noise level left by the erase head will be slightly higher than the virgin-tape level achieved by the bulk eraser. Possible sources of this excess noise include small changes in the erase field caused by the tape-to-head contact variations, the tape particle-to-particle magnetic variations, and the recording of Barkhausen noise from the erase core. The record head biasing field also produces similar increases in the noise level. The excess noise perceived by a listener due to the erase and record heads may rise as high as 6 dB above the virgin-tape noise floor.

28.3.8 Head Degaussing (Demagnetizing)

Early tape recorders used permanent magnets rather than an ac high-frequency signal to bias the tape so that small signals could be recorded without high distortion. These fixed magnetic fields produced a very high background noise level that severely limited the SNR of the taped recording. The introduction of ac bias upgraded the tape recorder from a voice-grade recording instrument to a true high fidelity recorder for music.

Modern recorders all use ac bias, but occasionally the background noise on a tape will be well above the normal level. The culprit is usually a permanently magnetized head, guide, or capstan that is acting like one of the old biasing magnets. The problem is most commonly created by touching a magnetized tool such as a screwdriver or razor blade to a component in the tape path. On rare occasions a faulty electronic circuit will create a dc current in one of the heads, leaving a residual magnetic field. (Loud clicks or thumps may be symptoms of dc currents.)

Since there are no commonly available instruments which can detect the very small magnetic fields which will result in noise, the best strategy is to frequently demagnetize all magnetic components in the tape path with a head degausser, Fig. 28-29.

The head degausser in Fig. 28-29 is an electromagnet with an extended core. The extension probe conducts an alternating magnetic flux generated in the coil to the tip of the probe. The probe is passed close to the magnetic components on the tape deck so that the alternating flux can flood the components. The actual demagnetizing occurs as the probe is slowly withdrawn from the component, creating the gradually decreasing alternating magnetic field mentioned previously in the discussion of bulk degaussers and erase heads.

Caution! Before using a head degausser, always verify that the tip of the probe is covered by a soft material that will not scratch the face of the magnetic heads. If necessary, wrap the tip with vinyl electrical tape or a similar tape.

Degauss the heads and other steel tape-guiding parts with a commercial-grade head degausser as follows:

1. Although a typical head degausser will not disturb a recorded tape that is more than a few inches from the degausser, always remove all tapes from the vicinity of the transport prior to energizing the degausser.
2. Hold the degausser at least 1 ft from the tape transport when applying power to the degausser. The degausser will produce a large voltage in the playback and record heads, which will probably not

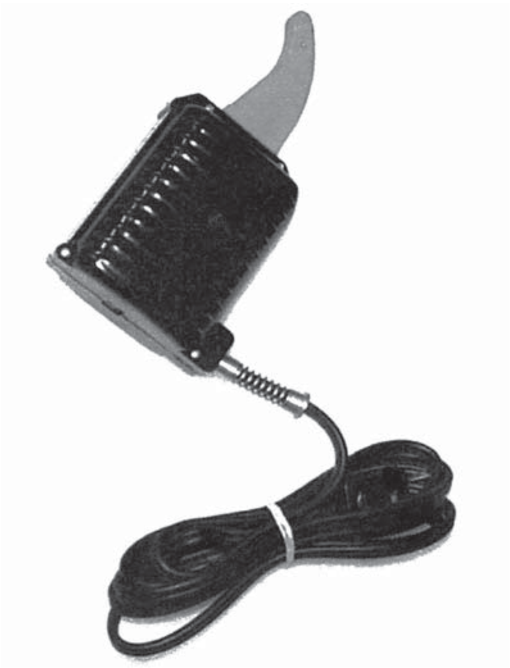


Figure 28-29. A head degausser.

damage the respective electronics but will certainly "peg" any analog meters in the circuit. Turn off the power to the recorder before using the degausser.

3. Move the degausser slowly and smoothly from bottom to top along the gap line of each head, moving at a rate of approximately 1 _ 8 in/s (3 mm/s). At the top of the head, smoothly withdraw the degausser 6 inches (15 cm) and then move smoothly to the next item to be degaussed.
4. To be safe, move the degausser at least 3 ft (1 m) away from the transport before disconnecting the power from the degausser.
5. Multiple degaussing passes on a component do not improve the quality of the results. A single slow, smooth pass is adequate.

The rapid collapse of the magnetic degaussing field at turnoff can easily undo all of the benefits of degaussing if the degausser has not been pulled away sufficiently. (For this reason, avoid degaussers that have momentary power switches that might be accidentally released in the middle of the degaussing routine.)

28.3.9 Tape Components

Modern magnetic tape consists of a powder of very small magnetic particles, which has been glued to one surface of a plastic substrate or base film. The backside of the substrate is coated with a very thin layer of carbon particles to improve winding characteristics and to reduce the buildup of static electricity.

28.3.9.1 Base Films

Although several base film materials were used in the past, including paper and acetate film, virtually all tape manufactured today uses polyester film (polyethylene terephthalate) such as Dupont's Mylar ™. Polyester is not only extremely strong and tear-resistant, but it is also relatively stable with respect to changes in temperature and humidity.

Depending on the intended application, the nominal base film thickness ranges from 1.4 mils (35 µm) for heavy-duty professional tapes down to a scant 0.25 mil (6.25 µm) for a C-120 cassette. To achieve reliable performance with these very thin films, the film must be not only very thin but also uniform in thickness from end to end and from edge to edge.

To enhance the strength of the thin base films used for cassettes, the polyester is prestretched or Tensilized. Although Tensilized tapes are more resistant to stretching than normal tapes, residual stresses that result from the Tensilizing process can produce physical distortion of the tape. For thin, narrow tapes these distortions are satisfactorily flattened out at the record and playback heads. The thicker, wider tapes used for professional formats, which would manifest severe contact problems due to these distortions, are considered to be strong enough without Tensilizing to provide adequate performance.

28.3.9.2 Binders

The glue or binder that holds the magnetic particles to the base film is a necessary evil that makes no active contribution to the magnetic performance of the tape. The use of new high-strength binders containing urethanes has improved both the durability and the recording characteristics of recent tapes.

The magnetic characteristics of the magnetic particles never wear out. The particles can be recorded and/or reproduced an unlimited number of times without any performance degradation.

The useful life of the tape is determined by three factors—the inherent strength of the tape, the amount of physical wear caused by the tape transport, and the performance required by the application. A typical test to measure the life of the tape would consist of many repetitive cycles on the intended transport while monitoring the gradual (hopefully) drop in playback level at the shortest wavelength of interest. When these losses exceed the application's requirements, the tape is worn out.

Some specialized audio transports designed for repetitive playback are capable of making over a quarter of a million passes on a tape. On the other hand, a poorly maintained studio recorder can destroy a master tape in ten passes or less! In general, if the abrasive forces exerted by the transport on the tape are well below the inherent strength of the binder, the tape will last virtually indefinitely. Any increase in the abrasive force due to dirty contact surfaces, excessive tape tension, or poorly designed tape guiding will accelerate the wear.

A very rapid catastrophic failure will occur once the abrasion force becomes sufficient to build up a small clump of debris on a contact surface. The friction between the debris and the tape surface is very high due to both the similarity of materials and the high pressure exerted by the tip of the clump as it pushes on the tape. The binder is overwhelmed, causing the clump to grow rapidly to the point at which the tape will show an obvious scratch or crease. If this situation should arise, the source of the problem should be corrected, and a copy of the damaged tape should be used for subsequent work.

From the magnetic performance standpoint, the combination of smoother magnetic particles and newer binders has enabled the tape manufacturers to use a smaller quantity of binder material to affix the magnetic particles. The ratio of useful particles to the magnetically inert binder rose from approximately 40% by volume for typical mastering tapes in 1970 to approximately 60% in 1980 with virtually no improvement since then. This improved magnetic density yields a higher maximum output for a given particle type and coating thickness.

28.3.9.3 Magnetic Particles

The ultimate performance of a tape recorder is determined not by the tape drive, heads, or electronics, but rather by the physical and magnetic characteristics of the magnetic particles of the tape. If basic performance parameters such as maximum output levels, noise, and distortion are truly determined only by the tape, the recorder is said to be tape limited. As a practical rule of thumb, if the noise and distortion products of the recorder are at least 10 dB lower than the products produced by the tape, the overall performance of the machine and tape will be within 0.5 dB of the theoretical levels of the tape alone.

Of primary importance in magnetic recording is the ability of each magnetic tape particle to assume and retain a magnetic pattern. These particles are chosen for their ability to maintain a magnetic field along one preferred direction or axis, permitting alignment of the particles for maximum performance. The amount of preferred orientation or anisotropy in the material depends on the nature and crystalline structure of the particles.

The shape of the particles determines the degree of physical alignment that can be achieved during the coating process. Smooth cylindrical or spherical particles that have no jagged edges or branches can be densely packed, yielding maximum output level.

The size of the particles is determined by the crystalline structure of each material. The residual noise of the tape decreases as the particles become smaller. Small particles with high anisotropy are therefore most desirable. Typical iron oxide magnetic particles for recording tape are cigar-shaped particles with a length-to-width ratio in the range of 4:1 to 8:1.

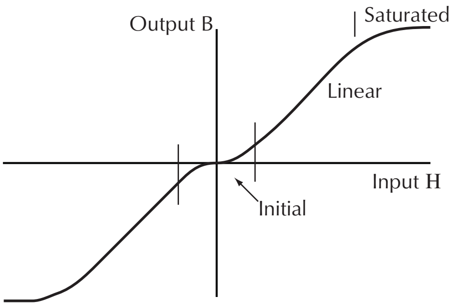
The newest recording products are abandoning particulate coatings in favor of thin layers that are plated or evaporated onto the surface of the plastic. These very thin layers of high coercivity materials are ideal for very short video wavelengths or very high digital bit densities. The new technology brings with it a whole new set of problems such as coating durability and how to include adequate lubrication in the metallic coating.

Coercivity. The *coercivity* is a measure of the magnetic force required to cause the tape particles to change magnetic polarity. High coercivity particles are more difficult to bias, record, and erase. On the beneficial side, they are also better able to resist external influences due to neighboring particles after recording, reducing the smearing of short-wavelength signals during storage.

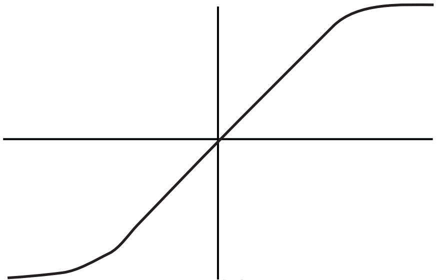
Retentivity and Remanence. If the coercivity is considered to be the input drive, then the *retentivity* and *remanence* are the output of magnetism left in the tape. Retentivity measures the maximum output per unit volume of coating cross section; remanence (remanent flux), which is the output per ¼ inch of tape width, varies not only with retentivity, but also with coating thickness. Remanence specifications should be used to compare the maximum long-wavelength outputs of different tape types.

28.3.10 Magnetic Performance Curves

The input-output relationship for typical magnetic materials is very nonlinear. As shown in Fig. 28-30A, the magnetization characteristic curve can be broken into three zones. For low excitation levels, the initial output is very small and nonlinear. As the excitation increases, a fairly linear region is encountered, which produces low distortions. As the level continues to increase, the magnetic particles finally become fully magnetized or saturated. Further increase at the input yields no more magnetization in the material.



A. Without bias.

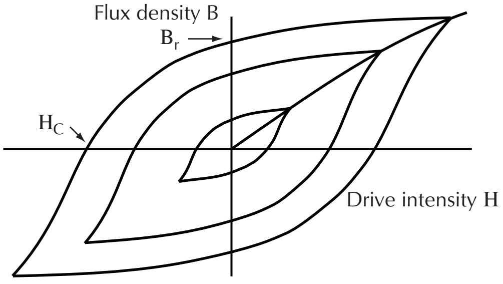


B. With bias.

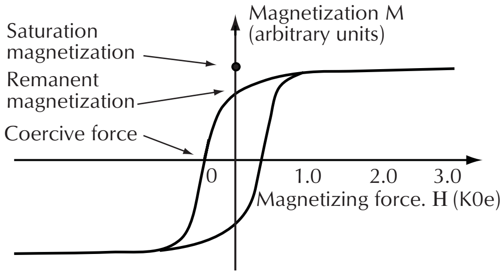
Figure 28-30. Tape transfer characteristics.

The nonlinear initial region must be avoided in audio recording if low distortion is to be achieved. The high-frequency bias signal provides enough excitation to jolt the magnetic particles into an active state. Optimizing the bias level yields the much more linear transfer characteristic of Fig. 28-30B.

Another representation of the magnetic characteristics is given by the B–H curves of the tape, as shown in Fig. 28-31. The curves show the amount of magnetic flux density created within the magnetic material by a cyclically varying intensity of applied magnetic excitation. Since the particles store part of the magnetic field, the path for increasing excitation differs from the decay path for decreasing excitation.



A. Varying levels of magnetization.



B. Critical parameters.

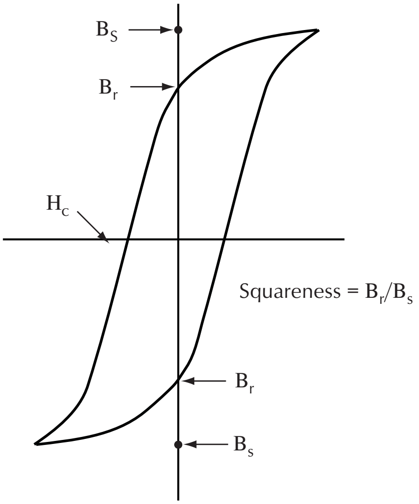
Figure 28-31. B-H curves.

Magnetic recording tapes are typically characterized by the coercivity and retentivity described previously. These points on the B–H curve for full saturation are indicated by H_c and B_r, respectively.

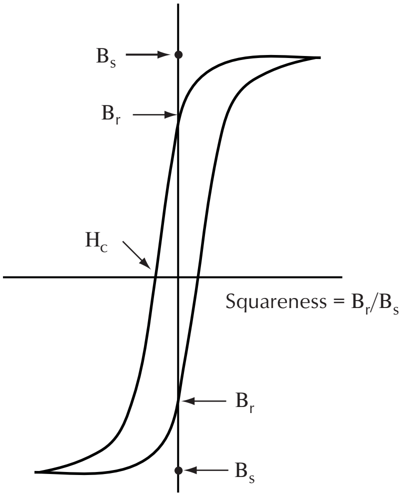
A figure of merit called squareness is commonly used to indicate the uniformity of the magnetic switching characteristics of magnetic coatings. As shown in Fig. 28-32, the squareness is the ratio of the remanent output value where the curve crosses the vertical axis to the saturated output. A perfect squareness of 1.0 would mean that every particle switched at exactly the same excitation level, yielding maximum output level and low distortion at high output levels.

The squareness ratio improves as more and more particles are aligned in parallel with the flux lines produced by the record head. The ideal case would be if all particles were exactly the same size with a perfect needle shape and all of the particles were stacked tightly like cordwood.

The early oxides had many branches or dendrites stick out of the sides of the needles. The dendrites interfered with the uniform packing of the particles, reducing the overall ratio of magnetic particles to binder. Later highly orientable particles (HOP) with reduced dendrites improved the packing factor. Additional work in coating techniques took advantage of the liquid flow of the coating during application to the base



A. Low squareness.



B. High squareness.

Figure 28-32. Squareness ratio.

film to align the particles, a technique known as *rheological orientation*.

As a result of these improvements, squareness ratios have increased dramatically. Over the past 30 years the squareness has improved from 0.8 to better than 0.9 for the best current audio tapes. This improvement translates to more available output and less harmonic distortion from the tape without requiring any increase in bias or record signal.

Table 28-5 summarizes the characteristics of several of the particles used for magnetic tapes.

Table 28-5. Characteristics of Tape Particles

		Size	Coercivity	Retentivity	Squareness
Fe_2O_3 (Gamma ferric oxide)	Low Noise	0.3 µm	350 Oe	1500 G	0.75–0.9
	Normal	0.7 µm	280 Oe	1100 G	0.75
Co-doped Fe_2O_3		0.3 µm	650–700 Oe	1300 G	0.75
CrO_2		0.6 µm	600 Oe	1500 G	0.9
Metal particle		0.2 µm	1100 Oe	3500 G	0.8

28.3.11 Magnetic Tape Specifications

The performance of a magnetic tape involves many parameters such as maximum output level, distortion, noise, print through, and frequency response. As a result, the data sheet that characterizes this performance must include many operating characteristics. The user must be very careful, however, to determine the test conditions under which the data is derived, including record head gap length, tape speed, operating level, and equalization.

One form for presenting this data is shown in Table 28-6. The data entries are measured for one specific recommended bias setting. Some of the values, such as sensitivity at long and short wavelengths, are comparisons to the performance of a standardized reference tape. The notes contain important information defining the test conditions used to derive the data.

Table 28-6. Tabular Tape Specifications

	Unit	Typical Values	Test Notes
I. Electromagnetic Properties			
Recommended Bias Setting	dB	3.0	1
Sensitivity at 1 kHz (81 kHz)	dB	0.8	2
Sensitivity at 10 kHz (Sl 0 kHz)	dB	1.1	2
10 kHz Saturated Output (SAT, 0 kHz)	dB	18.5	13
Third Harmonic Distortion at Reference Level (THD)	%	0.06	3
Output Level at 3% Third Harmonic			
Distortion (1 kHz) (MOLL kHz)	dB	17.5	4
Weighted SNR			
a. related to reference level	dB	–58.0	5
b. related to output level at 3% third harmonic distortion	dB	–75.5	5
Modulation Noise Ratio	dB	–73.0	6
Print through	dB	–58.0	7

Table 28-6. Tabular Tape Specifications (Continued)

	Unit	Typical Values	Test Notes
II. Magnetic Properties			
Coercivity (Hci)	Oe (kA/m)	350 (28)	8
Retentivity (Brs)	Gs (mT)	1500 (150)	8
III. Physical Properties			
Thickeners:			
Oxide Coating	mils	0.690	9
Backcoating	mils	0.040	9
Base	mils	1.400	9
Total	mils	2.130	9
Standard Widths:			
¼ inch	inches	0.246	
½ inch	inches	0.496	
2 inches	inches	1.996	
Width Tolerance:			
¼ inch	inches	+0.001	
½ inch	inches	+0.002	
2 inches	inches	–0.000	
Tensile:			
Yield Strength	lbs/qtr in	5.8	10
Breaking Strength	lbs/qtr in	11.6	11
Backcoating Resistivity	ohms/sq	5×10^4	12
IV. Measuring Conditions			
Tape Speed	in/s	15	
Reference Level	nWb/m	320	
Record Head: Gap Length	mils	0.50	
Track Width	mils	70	
Reproduce Head: Gap Length	mils	0.25	
Track Width	mils	70	
Reproduce Equalization	Ps	50 +3180	
Record Equalization		none	

Test Notes

1. Recommended bias setting is determined by adjusting the bias current for maximum sensitivity at 10 kHz andthen increasing the bias until the sensitivity changes by 3.0 dB. The adjustments made with a playback reference approximately 20 dB below reference level. The recommended bias setting corresponds to low third harmonic distortion and high output at 1 kHz.
2. Sensitivity is a measure of the output level compared to a standard reference tape A342D, when the recording is made at a constant input voltage approximately 20 dB below reference level and at the recommended bias setting.
3. Third harmonic distortion is the ratio between the level of the third order harmonic and the fundamental frequency (1 kHz) expressed in percent when recorded at reference level and at the recommended bias setting.

Table 28-6. Tabular Tape Specifications (Continued)

Unit Typical Values Test Notes

4. Output level at 3% third harmonic distortion is a measure of the output level capabilities of a tape at 1 kHz when recorded at 3% third harmonic distortion and at the recommended bias setting.
5. Weighted signal-to-noise ratio is defined as the ratio in dB between the 1 kHz output at reference level or at 3% third harmonic distortion and the ASA weighted (NAB standard) noise level. The noise measurement is made with the recommended bias and without input signal.
6. Modulation noise ratio is defined as the difference in amplitude between a 1.0 kHz signal level and its noise skirt at 800 Hz with a bandwidth of 10 Hz. The recording is made at reference level and the recommended bias.
7. Print through is the level of the accidental printing effect due to a signal recorded on an adjacent layer of tape. The printing signal is recorded at 1 kHz at reference level and the tape is held at 70° F for 24 hours.
8. Coercivity is the magnetic field required to reduce the magnetization of a saturated magnetic specimen to zero. The coercivity is a direct measure of the bias current requirement of a tape. Retentivity is the maximum remanent magnetization possible in a magnetic material. The long wavelength saturated output is directly proportional to the retentivity. Coercivity and retentivity values are obtained from a 60 Hz B-H loop tester with 1000 Oersted field calibrated to that maintained by the National Bureau of Standards.
9. Thickness measurements are made on Standard Gauges, 8000 Series, Smart Box.
10. Yield strength is defined as the force that produces 3% elongation of the samples. The measurement is made on an Instron tensile tester at a jaw separation of 5 inches and a crosshead speed of 2 inches per minute.
11. Breaking strength is the ultimate tensile strength indicating the force at which the tape breaks and is measured on an Instron tensile tester at a jaw separation of 5 inches and a crosshead speed of 2 inches per minute.
12. Backcoating resistivity relates to the tendency of magnetic tape to retain static charge. A resistivity value of 5×10^4 ohms per square is sufficiently low to prevent static buildup which might result in tape damage on high-speed bin loop duplicating systems or in normal use at low humidity conditions.
13. See bias curves.

Specifications are subject to change without notice.

In contrast, the graphical data in Fig. 28-33 depicts how the various values change as the bias value is adjusted over a range of 16 dB. All values are absolute values without any comparisons to a reference tape. The parameters of the recorder used for testing are shown above the graph.

The bias point recommended by the tape's manufacturer is the 0 dB value on the bottom scale. This value is a compromise value determined by simultaneously

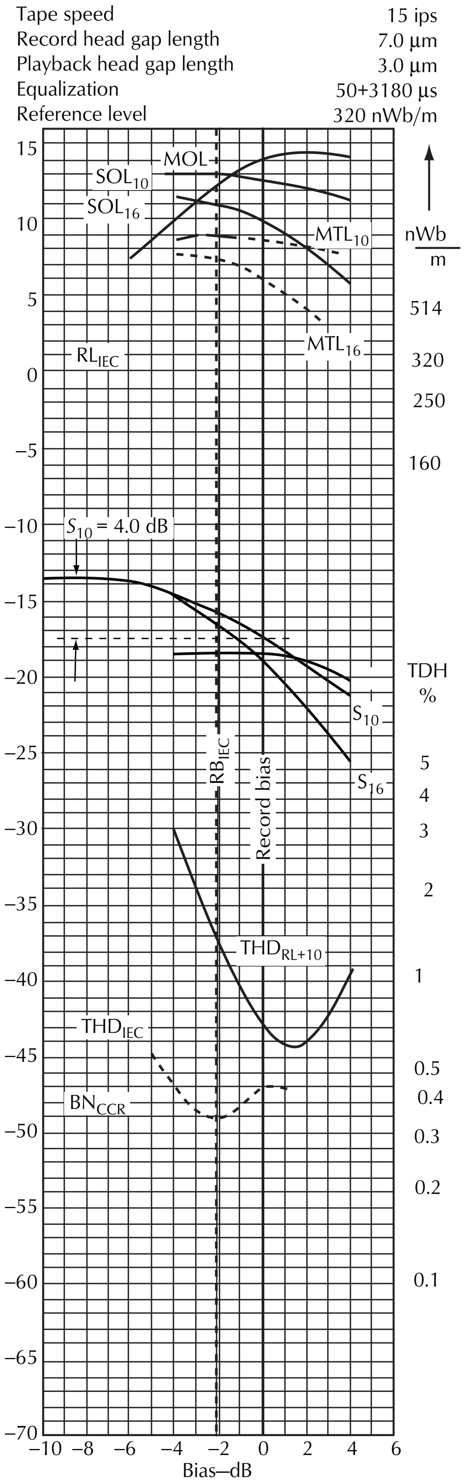


Figure 28-33. Tape performance data in graphical form.

evaluating the distortion, noise, and maximum output levels for each bias setting.

There are two common methods used for setting the bias level. One technique is to adjust the bias while recording a long wavelength such as 1 kHz. The bias is increased until the recorded signal peaks. The bias level is then further increased until the recorded signal drops by 0.5 dB.

A second technique is to use a short wavelength, typically 1.5 mils, and adjust for a significant amount of overbias. The bias is increased until the recorded signal peaks. The bias is then further increased until the recorded signal decreases from peak by several dB.

How do these two techniques compare? Find the sensitivity curves S_1 and S_{10} near the center of the graph. These curves show how the 1 kHz and 10 kHz signals will change in level as the bias is increased. Note that the S_1 curve is very flat, changing only ¼ dB from peak over a bias range of 5 dB. In comparison, the S_{10} curve is falling at a rate of approximately 1 dB/dB of bias increase.

The flat shape of the S_1 curve provides very little signal drop for a rather large bias change. A 0.1 dB error in the signal level adjustment, perhaps due to a sticky meter, may change the 10 kHz sensitivity by 2 dB or 3 dB. This error would require an additional record equalization boost of 2–3 dB to correct the overall response.

In contrast, the rapid signal level change when using a 10 kHz signal gives a much more precise adjustment and better uniformity from track to track. It is clear that both techniques are trying to achieve the same adjustment, but the short wavelength technique offers much better resolution.

This technique can be a trap for those who don't understand what is actually happening. The S_1 and S_{10} curves are really curves for the specific wavelengths of 15 mils and 1.5 mils, respectively. If the tape speed is doubled, these curves will now represent performance at 2 kHz and 20 kHz. The 15 in/s overbias specifications for 10 kHz must not be used at any other speed! For example, at 30 in/s the S_{10} curve of the example tape has a downward slope of only 0.5 dB/dB of bias. Why? Because the wavelength is 3 mils, not the 1.5 mils of the previous example at 15 in/s. The manufacturer recommends only 1.5 dB of overbias at 10 kHz and 15 in/s. It is important to use the same wavelength at all speeds by shifting the test frequency to 20 kHz or 5 kHz at 30 in/s and 7.5 in/s, respectively.

As mentioned previously, the test data is very dependent on the characteristics of the recorder used during the testing. In particular, the shape of the S_{10} curve varies greatly with changes in the record head gap length. Table 28-7 illustrates how this gap length affects the recommended amount of overbias.

Table 28-7. Short-Wavelength Dependence on Record Gap Length

Record Gap Length	10 kHz Overbias @15 in/s
1.0 mil	1.0 dB
0.5 mil	2.5 dB
0.25 mil	3.0 dB

28.3.12 Problems with Older Tapes

The archives of American tape recordings contain tapes that are up to 50 years old. Unfortunately, many of these tapes have problems that could easily damage or destroy their recordings. Some of these problems can be corrected, but others are irreparable.

28.3.12.1 Adhesion and Peeling Oxide

Adhesion is the binding force that firmly holds the oxide layer onto the plastic substrate. Two simple tests can be used to evaluate the strength of the adhesion—the Scotch tape test and the sharp edge test.

The Scotch tape test tries to rip the oxide from the plastic substrate by brute force. Start with a strip of Scotch Brand Magic Mending tape several inches long. Adhere about 3 inches of the sticky tape to the oxide surface of the recording tape. Rub the joint to assure complete binding of the tapes. The test is to peel back the sticky tape with a quick jerk parallel to the tape. If the sticky tape comes off cleanly, the adhesion is good. If the oxide layer delaminates and peels off with the sticky tape, the adhesion is poor.

The second adhesion test utilizes a blunt edge to create a very sharp bend in the tape. Find a sharp perpendicular edge on a desk or a piece of plastic ruler that has no rounding radius. Place the backside of the tape sample against the sharp edge. Pull on the ends of the tape to establish a firm tension in the tape. While maintaining the tension, drag the tape over the edge in a 90° bend. If the oxide does not loosen from the backing, the tape passes the test. A poorly adhered tape may suffer complete delamination of the oxide, with a solid band of oxide peeling off and shooting away from the backing.

28.3.12.2 Brittleness

The polyester base films and urethane binders of modern tapes remain flexible under all normal circumstances. The base films and oxide layers of earlier tapes, however, could become brittle. Plasticizers were included in the binders and the acetate backing to provide flexibility in the tape. Unfortunately, these plasticizers can harden with age, causing the tape to become brittle. Harsh storage conditions can accelerate the breakdown of the plasticizers.

Brittleness cannot be reversed. The only remedy is to use a tape transport that is extremely gentle on the tape. Choose a transport with dynamic braking rather than harsh mechanical brakes. A transport with constant tape tension can be set for the lowest practical tape tension. Some decks also feature gentle start capability that ramps the capstan up to speed to smoothly accelerate the tape rather than just slamming a pinch roller onto a running capstan.

28.3.12.3 Splice Failures

In the early days, standard Scotch Brand cellophane tape was the only splicing tape. Later on, splicing tapes such as Scotch #41 and #620 were developed with improved characteristics. Although these tapes were fine for day-to-day operations, they have not survived the test of 50 years of storage. For example, the adhesive of both cellophane tape and splicing tape can ooze out and stick to adjacent layers of tape. A common remedy was to apply talcum powder to the sticky oozed adhesive to avoid layer-to-layer adhesion.

With even more time the adhesive can dry out completely, causing the splice to fail. In this case the only remedy is to replace the splicing tape with new tape. The newest splicing tapes, such as blue #67, replace the original latex adhesives with synthetic adhesives that do not ooze or dry out.

The tape operator must be watchful for two problems created by bad splices. First is layer-to-layer adhesion that can produce strong tugs that break older acetate tapes. Second is tape separation at a failed splice. DO NOT run the tape through the recorder at high speed if you suspect either problem may exist. If the tape tugs or separates at high speed, the loose end may be slapped around and broken off before you can stop the spinning reels.

If the tape is old, wind the tape slowly and carefully to examine each splice. Re-splice all splices if there is any hint that the splices may separate. Do not try to remove any of the old splicing tape adhesive that has dried out on the recording tape. Sticky adhesive residue that could bond to adjacent layers, on the other hand, must be removed.

28.3.12.4 Print Through

The energy required to activate a particle to switch magnetic states depends on the size of the particle, with the overall characteristics of a magnetic tape being determined by the average size and characteristics of many particles in the coating. A more detailed analysis of the particles would yield a distribution of sizes, as shown in Fig. 28-34. Although the majority of the particles cluster around the average value, a small portion of the particles are either much smaller or much larger than the average. The small particles give rise to spontaneous recording as print through; the large particles produce noise bursts.

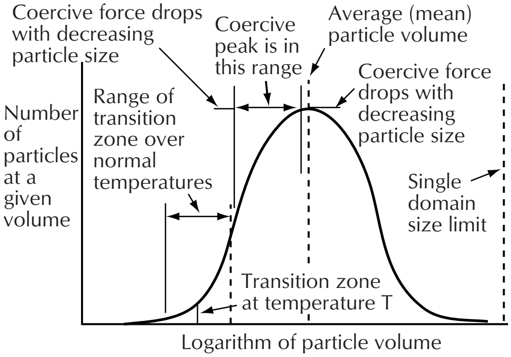

Coercive force drops with decreasing particle size
Coercive peak is in this range
Average (mean) particle volume
Coercive force drops with decreasing particle size
Number of particles at a given volume
Range of transition zone over normal temperatures
Single domain size limit
Transition zone at temperature T
Logarithm of particle volume

Figure 28-34. Particle size distribution.

The small particles require so little activation energy to assume a new magnetization state that even the thermal energy of the particles may provide enough bias to cause the particles to be recorded by the stray magnetic fields due to adjacent layers of recorded material. This spontaneous recording is most evident as pre- or post- echo at the beginning and end of a recording. The strength of the print through image depends on both the percentage of thermal idiots in the coating and the ratio of remanence to coercivity of the tape. The remanence measures the driving force of the signals trying to print through. The coercivity, on the other hand, is the stubbornness of the particles to resist this imprinting. The effective coercivity of the small particles is diminished because the domain size is sub-optimal, rendering the small particles more susceptible to printing.

The milling process used to provide thorough mixing of the particles, binder, and additives prior to coating is a rather abusive process that can create thermal idiots by fracturing some of the desirable large particles into smaller, low-coercivity particles. Insufficient milling, on the other hand, provides an uneven particle dispersion that creates noise on the tape. The tape manufacturer must strike a compromise that yields both low noise and low print through.

Print through of a signal produces both pre-echoes and post-echoes. The pre-echoes are more troublesome in music, however, since the pre-echoes frequently occur in the quiet passages just before the loud note. The post-echoes, on the other hand, are frequently masked by the diminishing tail of the musical note and the room reverberation.

Fortunately, the print-through process does not produce equal amounts of pre- and post-echo, but unfortunately the more undesirable pre-print echo is the stronger. The vector magnetization components that arise during the recording process cause the levels of print on the outer adjacent tape layer to be several decibels higher than on the inner adjacent tape layer, as shown in Fig. 28-35. The more troublesome pre-print echoes on musical selections can therefore be minimized by storing the tape tails out to move the quiet lead-in to the inner layer. This will also bury the louder outer layer print through echo in the decaying signal at the end of the music.

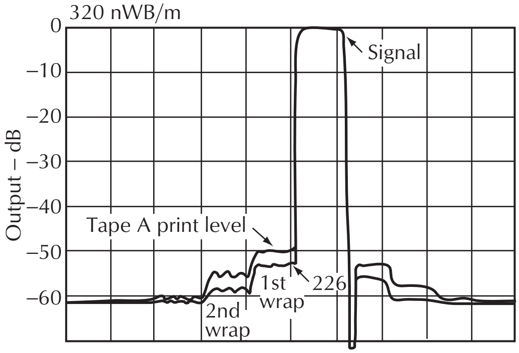


Figure 28-35. Pre- and post-echo print through. Courtesy 3M Co., Magnetic Audio/Video Products Div.

The use of nonmagnetic leader tape between selections is also helpful to eliminate pre-echo on selections that begin with a rapid attack. Be aware, however, that paper leader tape can contain a small amount of magnetic debris that will raise the noise level as the leader passes over the playback head.

The user can take several steps that will minimize the amount of print through. First, the use of thicker base films increases the spacing between layers.

Second, avoiding elevated temperatures and stray magnetic fields during use and storage will decrease the excitation of the thermal idiots. Third, exercising the tape by shuttling the tape from reel to reel several times will partially erase the printed particles. The flexing and rubbing of the tape produce enough activation energy to neutralize some of the printing. For this reason, never copy a stored master tape without exercise. The worst possible print through level exists on the very first pass of the tape. In some cases print through can drop as much as 4–6 dB with five shuttle cycles.

28.3.12.5 Sticky Shed and Tape Baking

As mentioned earlier, the goal is to attach a maximum number of perfectly stacked and oriented magnetic particles onto the surface of the backing material. Anything that interferes with this goal by displacing some of the magnetic particles, such as additives for lubrication, fungicides, and static charge reduction, degrade the performance of the tape. Most important in this category is the very binder that holds the particles in place. Every bit of binder displaces some of the useful magnetic particles.

The best choice is to use a very strong binder that can do the job with the minimum amount of glue, allowing space for more magnetic particles. The winner is the highly crosslinked thermoset polymers in use today. Starting around 1970, these binders with a high urethane content gave a big boost to tape performance. Unfortunately, long term experience with these tapes now shows that the binder can break down during storage. The symptoms are a buildup of residue on the head and guides and a tendency for the tape to stick to these residue buildups, in some cases actually dragging the tape to a stop. The popular name for this problem is sticky shed. The problem is usually discussed in terms of binder breakdown, but there appears to be a second major problem related to lubricant oozing that is also present.

Urethane Binder Breakdown. The urethane binder contains long polymer chains that provide the high strength of the binder. Water in the surrounding air enters the tape and breaks the long chains through a process known as *hydrolysis*.

As a result of the chemical breakdown of the long polymer chains, the binder is weakened enough for the surface of the tape to begin to rub off onto the stationary guides and heads. Depending on the design of the transport, this residue can clog the heads in just a few seconds. Machines with rotating guides and low tape tensions take longer to clog, but the damage to the tape is still intolerable.

Fortunately, the hydrolysis is somewhat reversible. Tapes can be *baked* at a moderate temperature to reverse the hydrolysis and restore strength to the binder. Although this may seem like a bit of witchcraft, thousands of baked rolls of archived tapes have proven the technique.

The electric oven must provide a well-controlled temperature of about 120–140°F (50–60°C). Large dehydrators or fruit dryers are popular because of their size and limited temperature range. Only an electric oven should be used. The oven should be preheated and checked for temperature stability with a high-resolution thermometer such as a candy thermometer. The tape, wound onto a metal reel, is placed into the oven horizontally with generous space above, below, and around the reel for air circulation. The tape is baked for 15–20 hours, and then allowed to cool to room temperature undisturbed in the oven.

The baking process creates a low-humidity environment that draws some of the excess water from the tape binder. The short polymer chains may recombine with their neighbors to produce a better bond, but the breakdown process is not fully reversed.

Lubricant Oozing. The second failure mechanism also involves the binder, but in this case the culprit is the oxide. A change of oxide particles also changes the chemistry needed to make a liquid binder that can:

1. Hold all the magnetic particles in suspension.
2. Be smoothly coated onto a polyester backing material.
3. Have the volatile byproducts evaporated in the drying ovens.

In the early 1970s Phizer introduced a new high-output oxide with excellent signal characteristics, but the particle required a reformulated binder with a low pH in order to meet the above requirements for a usable dispersion. This particle was utilized by 3M in the 226 family of tapes (226, 227, 806, 807, 808, 809) and by Ampex in the 456 family.

The new binder formula included a component that served primarily as a lubricant. Unfortunately, however, this lubricant would migrate to the surface of the tape and concentrate into a sticky residue.

The baking operation described in the previous section warms the concentrated lubricant enough to allow the lubricant to flow and be reabsorbed into the depths of the coating.

Since both types of sticky shed problems are treated by baking, most people who bake tapes don't know for certain which problem they are treating, and if the sticky shed is eliminated, they probably don't care.

How long before a baked tape begins to again exhibit sticky shed? Results will vary depending on the amount of degradation, the tape type and specific batch, the exact baking method, and the operating environment after baking. Reports vary from days to years. Certainly baking provides an adequate window for the tape to be transferred to another medium.

Is there any degradation due to the baking process? The most likely problem is print through caused by the elevated temperature. Print through is a time-dependent problem that peaks out at a maximum value after a long time. Heat accelerates the printing. However, stored tape probably has had enough time for the print through to be near the maximum value before baking. As a result, the additional print through caused by the baking may be negligible. Follow the exercise process described at the end of Section 28.3.12.4 to minimize the print through before copying the tape.

How can sticky shed be avoided? The rate of hydrolysis depends on the storage conditions. Archival storage at a temperature of 60°F (15°C) and relative humidity of 25% ± 5% is optimal, but few have the luxury of an environmental chamber, so store the tapes in a cool, dry location in the original package standing on edge.

Sticky shed may also produce layer-to-layer adhesion. If you strongly suspect sticky shed, bake the tape before trying any winding operations on a tape transport. This will avoid the total loss of recorded segments due to oxide being ripped from the tape's plastic backing during spooling.

28.3.12.6 Squealing Tape

One of the many ingredients in the coating recipe is a small amount of lubricant. Obviously, the tape cannot be too slippery or else the capstan couldn't maintain constant tape speed. Running the tape completely dry, on the other hand, can produce an audible squeal. The tape undergoes a "stick-slip" phenomenon on the stationary guides and head, creating a jerky motion at a high frequency. The irregular motion can even be measured with a scrape flutter meter.

The squeal results from the loss or failure of the original lubricant. The obvious solution is to replace the lubricant. A can of 10-W30 motor oil isn't appropriate, but another common household lubricant, WD-40, is recommended by Quantegy. Quantegy claims that WD-40 is cheap, available, inert to all recorder components, and a very good lubricant. Apply the oil sparingly by lightly wetting a lintless rag or swab with the oil and holding the applicator against the oxide side of the moving tape at the first guide after the supply reel. A bit of experimenting may be required to find the proper amount of oil that is required to eliminate the squeal without causing slippage and speed irregularities. When you are finished using the tape, prepare the tape for storage by passing the tape over a dry applicator in a medium speed fast wind mode. (Use extreme caution on tape transports that have elastomer coatings on the capstan and/or timing rollers. Lightly lubricate the tape while passing the tape directly from the supply reel to the takeup reel, and then dry the tape with a second pass over a dry applicator before threading the tape over the elastomer components.)

28.4 Analog Circuits and Systems

The transport mechanism, heads, and tape should combine to determine the basic performance limitations of a tape recorder. The analog electronic circuits of the recorder, on the other hand, should exceed the capabilities of the heads and tape in all respects so that only the heads and tape limit the quality of the final signal. In practical terms, this means that the SNR, frequency response, distortion, and head room of the electronics are comfortably better than the heads and any tape, including future improved tapes.

The block diagram of the signal electronics of a typical professional audio recorder is shown in Fig. 28-36. In terms of actual hardware, approximately 75% of the audio circuits of a modern professional audio recorder are devoted to operator interfacing and controls; the remaining 25% implement the basic functions of erasing, recording, and playback. Since the variation of features and technology used to implement the interfacing and control functions is too broad to be summarized herein, the following description covers only the latter basic functions.

28.4.1 Playback Amplifiers

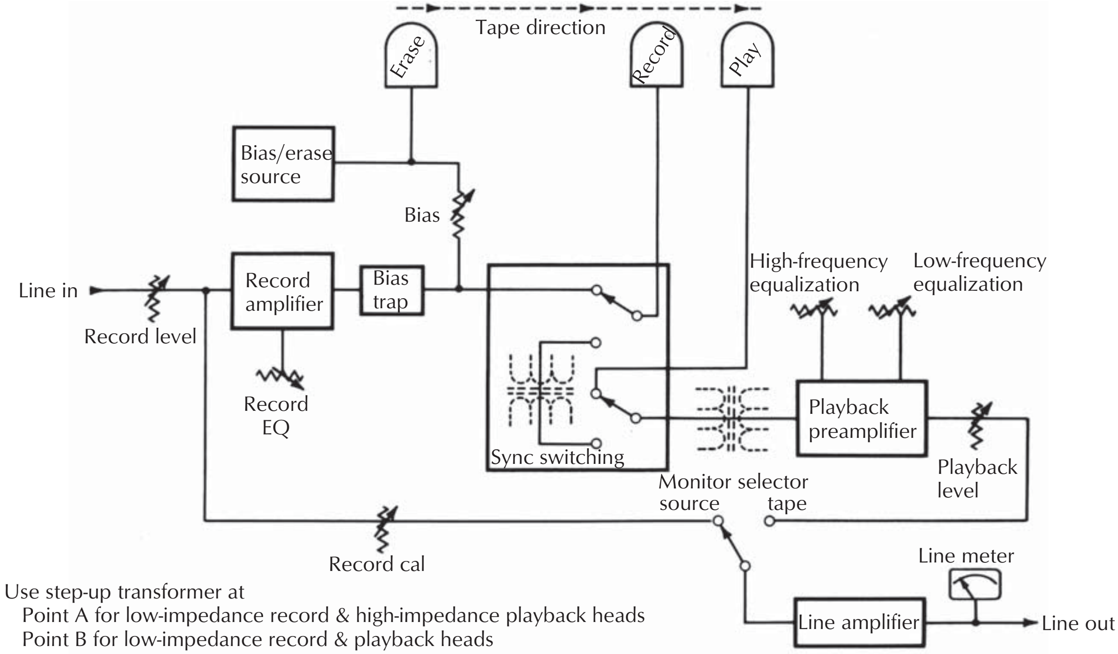
The amount of electrical power that can be generated by a magnetic tape passing over the face of a reproduce head is exceedingly small. The output voltage from the head for loud recorded passages will reach no more than a few millivolts, with quiet passages dropping into the microvolt region. This weak signal must be carefully boosted without the introduction of additional noise to a higher, more usable level by the first stage of the playback amplifier. Special low-noise amplifier circuits developed for this purpose provide at least 20 dB of gain so that subsequent amplifier stages will not be required to operate near their noise limits.

Since the reproduce head produces an output voltage that is related to the rate of change of the flux on the tape, $d\phi/dt$, the output voltage will rise at a rate of 6 dB/octave. A compensating circuit with a falling 6 dB/octave response, known as an integrator, is used in the playback amplifier to correct for this rise and give a voltage that is proportional to the value of flux sensed by the head.

When the effects of playback head resonance peaking and gap length, spacing, eddy current, and thickness losses are included, the output from the low-noise amplifier and integrator would follow the falling curve in Fig. 28-37 for 15 in/s (38 cm/s) operation, see Fig. 28-12 for details. This curve must be reshaped by the combined effect of the record and playback equalizers to yield a flat response. The method of partitioning this correction between the record and playback circuits is dictated by the equalization standard chosen by the operator. Since all users of a given equalization standard will be using the same partitioning, the recorded tapes will all be interchangeable.

Fig. 28-38 shows a simplified schematic of a typical operational amplifier type of a playback amplifier capable of the necessary playback corrections. The low-frequency-cut circuit is utilized in some NAB and cassette standards to achieve a decrease in low-frequency playback noise below 100 Hz at the expense of low-frequency headroom. A typical design would include additional ancillary components for amplifier biasing and stabilization.

With one common exception, the same type of circuitry is utilized in the sync/overdub mode to condition and amplify the playback signal from the record head. The exception is in the form of an added voltage-boosting transformer that is commonly necessary to get the signal above the noise floor of the low-noise input section. This problem arises from the low inductance and few turns of wire that are typically found in a record head. The record head must pass the audio plus the high-frequency bias signal; therefore, the inductance must be kept low enough to avoid self-resonance with the head cables at the bias frequency. When fewer turns are used to reduce the inductance, the output voltage goes down proportionately. In essence, these turns are restored in the transformer by a step-up turns ratio ranging from 3:1 to as high as 10:1.



Use step-up transformer at
Point A for low-impedance record & high-impedance playback heads
Point B for low-impedance record & playback heads

Figure 28-36. Tape recorder signal block diagram.

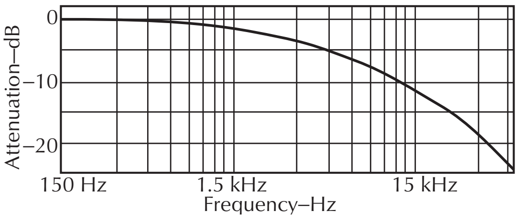


Figure 28-37. Unequalized reproduce head output.

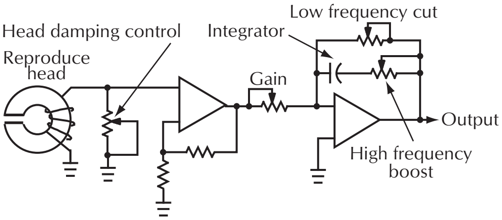
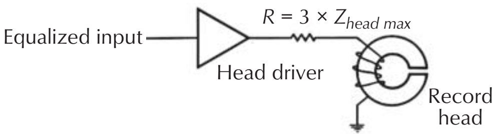
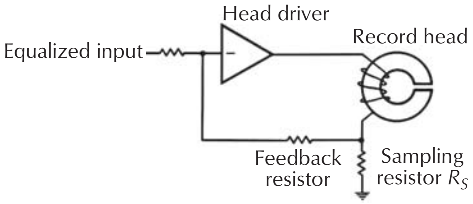


Figure 28-38. Simplified playback amplifier.



A. Combined series impedance of head and resistor remains constant.



B. Sampling resistor in series with return leg of the head.

Figure 28-39. Constant-current record head drivers.

28.4.2 Record Circuits

The primary task of the amplifier that drives the record head is to convert the input audio signal voltage into a proportional amount of current flowing in the windings of the record head. To accomplish this task, the head driver must overcome the rise in head impedance with frequency that is due to the inductance of the head. A common technique to achieve flat current response, as shown in Fig. 28-39A, is to insert a resistor in series with the head so that the combined series impedance of the resistor and the head remain relatively constant throughout the audio band. If the resistance is chosen to be two to three times the reactance of the head at the upper limit of the desired audio band, the desired constant current characteristics can be closely approximated.

The primary disadvantage of the series resistor is the loss of head room due to the extra signal drop across the resistor. This problem can be overcome with an active current feedback circuit that senses the current in the head through a small sampling resistor. Fig. 28-39B shows a sampling resistor R_s in series with the return leg of the head. The voltage generated across R_s by the current flowing in the head is fed back to the inverting input for comparison with the incoming audio signal. The high gain of the driver amplifier necessitates only a very small feedback signal, creating a negligible loss in head room at high frequencies.

The circuits of Fig. 28-39 oversimplify the task of driving the record head since no provisions are included for adding the high-frequency bias signal to the current in the head. A common method of adding the bias and audio signals is shown in Fig. 28-40. The audio driver is isolated from the bias source by a parallel trap tuned to the bias frequency so that the bias signal does not create nonlinearities within the audio driver. The high impedance of the trap at the bias frequency also reduces the loading effect of the audio source on the bias source.

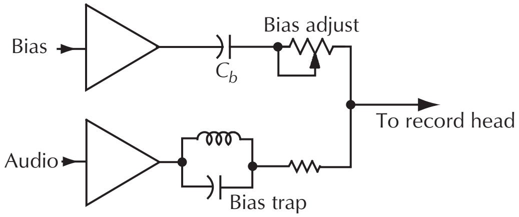


Figure 28-40. Audio and bias coupling to record head.

A similar isolation of the bias source is accomplished by the capacitor in series with the bias supply. Since the capacitor looks like a high impedance at audio frequencies, the loading effect of the bias supply on the audio source is minimized. At the higher frequencies of the bias signal, the reactance of the capacitor has dropped to a relatively low value that provides adequate coupling of the bias signal into the record head.

An alternate approach that eliminates some of the previously mentioned isolation requirements is shown in Fig. 28-41. In this case, the bias and the audio are added together at the input to a combination bias/audio head driver amplifier. If the amplifier has sufficient head room and very low distortion, the two signals can be amplified simultaneously by the same amplifier without any interference. The problem of coupling the output to the head for constant current drive must still be overcome, however, by including either a complex coupling network or an active feedback network.

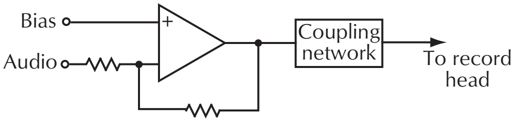
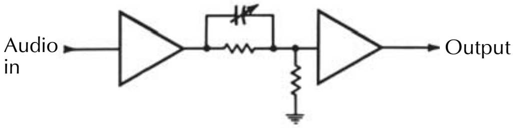


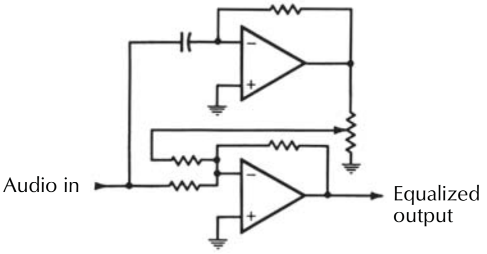
Figure 28-41. Active summer for bias and audio.

In addition to the head driver circuits, which correct for any response droop due to head inductance, the record amplifier must provide deliberate frequency response tailoring to match the desired equalization standards. The standards usually require an adjustable boost at 6 dB/octave beginning in the middle of the audio band, with lower tape speeds generally requiring greater boosts to overcome the increased tape thickness and self-erasure losses.

The needed boost is easily implemented by the resistance-capacitance circuit shown in Fig. 28-42A, but the use of a variable capacitor is inconvenient due to the limited range of capacitor adjustment and the awkward size and mounting of the capacitor. Newer designs, therefore, favor operational amplifier configurations that control the amount of boost with a potentiometer. One such circuit, shown in Fig. 28-42B, selectively adds the output of a differentiator circuit, which rises at 6 dB/octave, to the main signal path.



A. With variable capacitor.



B. With operational amplifier configuration.

Figure 28-42. High-frequency record boost circuits.

A secondary benefit of the differentiator circuit is the phase change introduced by the inverting characteristics of the differentiator amplifier. Unlike most of the loss-correction circuits of the signal path, which introduce signal delay at high frequencies, the inverted differentiator advances the high frequencies. The proper combination of advance and delay can provide less phase distortion in the signal, yielding improved transient response with less overshoot. A similar phase-correcting effect has been implemented in other designs by providing an all-pass, phase-shifting network in the reproduce amplifier.

The NAB and compact cassette equalization standards provide an additional record signal boost at low frequencies to overcome the hum and noise limitations of the reproduce heads and amplifiers. Typical circuits for this purpose are shown in Fig. 28-43. Both cases achieve a 6 dB/octave rise with decreasing frequency from 50 Hz or 100 Hz to below 20 Hz.

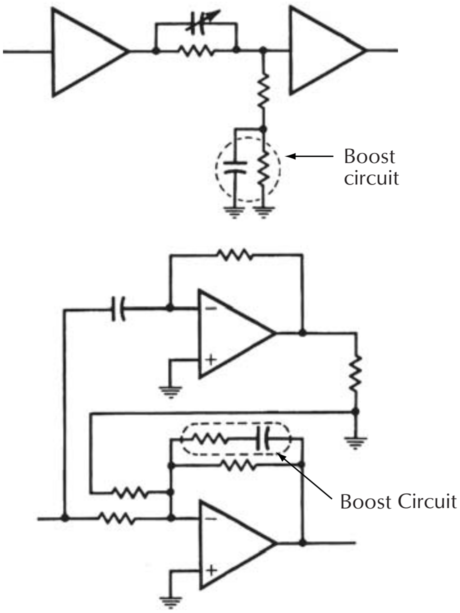


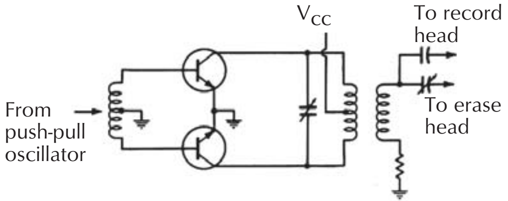
Figure 28-43. Low-frequency record boost circuits.

Abrupt changes in the bias and audio signals on the record head must be avoided whenever the record mode is entered or exited. Ramping circuits are employed for this purpose to control the buildup and decay of these signals. Typical methods include the use of analog switching elements such as bipolar-junction or field-effect transistors. The rate of switching of these elements is limited to a value that does not create abrupt transients but, at the same time, is quick enough to avoid annoying delays, overrecordings, or program holes.

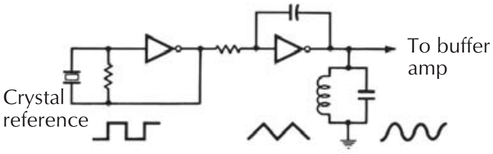
28.4.3 Bias and Erase Circuits

The high-frequency signals required for biasing and erasing all tracks of the tape are derived from a single master oscillator so that no interference or beating of multiple oscillators will occur. Older designs generally employ a tuned push-pull multivibrator oscillator; newer designs favor crystal-stabilized oscillators utilizing digital circuitry. Several designs have used separate bias and erase frequencies, with the erase circuit running at one-third the bias frequency to minimize the power dissipation on the erase head.

In all cases, the primary consideration is purity of the bias and erase current waveforms. Any even-order harmonics, including dc, second harmonics, fourth harmonics, and so on, will create a detrimental rise in the background tape noise, reducing the available SNR for the recorder. Older designs, such as Fig. 28-44A, relied heavily on push-pull circuits with balancing transformers to minimize these even-order components. Newer designs, such as Fig. 28-44B, favor filtering and feedback control to reduce unwanted components. The divide-by-two flip-flop eliminates any even-order distortion in the oscillator waveform.



A. Push-pull circuit with balancing transformers.



B. Filtering and feedback control.

Figure 28-44. Typical bias and erase sources.

The erase head is typically coupled to the erase source with an adjustable series resonating capacitor to minimize the voltage required from the driver and to filter out even-order components. A current sampling resistor is frequently provided in the ground leg of the erase head circuit so that the amplitude of the erase current can be conveniently monitored.

28.4.4 Noise Reduction Systems

The SNR of an analog audio recorder is usually taken as the difference between the residual biased tape noise level and the level which produces 3% third harmonic distortion at 1 kHz. In the ideal case this ratio is limited by the tape speed, track width, and tape type. Once these parameters are set, the maximum SNR is determined.

Direct analog noise reduction systems rely on the masking effect of human hearing. If both a background noise and a louder desired signal exist within the same frequency band, the noise will be masked by the desired signal. If, on the other hand, the noise and signals are in different parts of the audio spectrum, such as a bass guitar and high-frequency tape hiss, the noise will not be masked. The perceived noise can be reduced if the SNR is compromised during masking situations so that unmasked noise can be reduced. This requires dynamic change of the gain or transfer function of the system depending on the program content.

Dolby™ and dbx™ noise reduction systems are examples of amplitude-only encode/decode systems. Both systems modify the amplitude of the signal to squeeze the dynamic range of the input signal into a smaller dynamic range that will avoid the noise and distortion limitations of the recording tape. The fidelity of these compander (compression/expander) systems is limited not only by the tracking of the encode and decode circuits, but also by the nature of the errors that are generated by noise, nonlinearities, and frequency response anomalies introduced by the record/playback cycle of the tape recorder. These parasitic errors can cause dynamic mistracking that will create distortions of dynamic signals that may not be evident during sine-wave testing.

The Dolby systems process the low-level signals by boosting them during recording and then attenuating them on playback. The original professional Dolby system (Dolby A) subdivides the audio spectrum into bands that are processed individually to optimize the masking effect.

A later development, the Dolby SR system also adds adaptive filters that change their cutoff frequencies as the signal content varies. When the program material includes information at high frequencies, the filter opens up to full bandwidth. If no high-frequency content is present, the cutoff frequency of the filter slides down to match the program material. The filter must be "intelligent" enough to distinguish between desired audio signals and unwanted noise.

The Dolby SR system was quickly embraced by the music and film markets as a method of raising analog recorder performance to a level that rivals digital recorders.

DBX is also a "compander" system, but instead of dividing the spectrum into bands and acting on each band differently, DBX acted on the true RMS value of the signal. This resulted in SNR improvements that were significantly better than Dolby A, but many users felt that the DBX system was more prone to audible artifacts of the process.

28.4.5 Sync Operation

Multitrack recording requires that artists be able to listen to the previously recorded tracks while simultaneously adding their new performance in synchronism with the prior tracks. Analog recorders accomplish this by using some tracks of the record head as playback sources while simultaneously recording on other tracks of the same head.

28.5 Digital Magnetic Recording

The information in this section on digital magnetic recording is presented as an attempt to document a bit of audio history. It is very likely that by the time the next edition is published, all recording will be performed in RAM or other solid state memory, and magnetic recording both analog and digital will be of historical interest only. As of this publication there is a clear trend toward digital audio workstations as being the standard recording devices, with the audio stored on hard drives. Tape-based systems are virtually gone. Currently, computers exist which utilize flash memory, and as such have no moving parts at all. At the moment these computers are expensive and have relatively low capacity compared to conventional hard-disk-based computers, but they are clearly the wave of the future.

28.5.1 Longitudinal Digital Tape Transports

Longitudinal digital tape recorders came in many varieties, the most successful were the DASH and ProDigi formats. DASH is an acronym for Digital Audio Stationary Head. The tape transports for these recorders were very similar to any high-quality analog mastering machine, but the very high density of very narrow tracks required extremely accurate tape guiding and head placement. For example, the DASH format with 52 tracks across a ½ inch tape specified head height to 0.6 mils (0.015 mm) and a guide placement to 0.2 mils (0.005 mm).

Tape speeds for multitrack DASH and ProDigi machines was 30 in/s for normal 48 kHz sampling of 16 bit data, but the DASH high resolution (HR) upgrade boosted the speed to 45 in/s for recording 24 bit data. The tape speed was servo controlled by the capstan to exactly match the sampling rate of the data recorded on the tape. The sampling rate could be varied from nominal by up to ±7% for varispeed operation.

Both formats included tape cleaning devices to remove loose debris from the surface of the tape. Loose particles were wiped from the oxide surface by passing the tape across a post covered with lintless fabric tape. A clock motor mechanism slowly advanced the fabric to refresh the wiping surface.

The DASH and ProDigi formats used conventional reels of tape that were mounted on the reeling spindle and threaded through the machine by hand. Aside from a few types of reel-to-reel digital instrumentation recorders, all other modern longitudinal digital tape formats utilize self-threading tapes that are permanently enclosed in a cartridge or cassette.

28.5.1.1 Signal Flow

A 48 channel DASH digital recorder contains more electronic circuits in one channel than all the audio electronics in an entire 24 track analog recorder, Fig. 28-45.

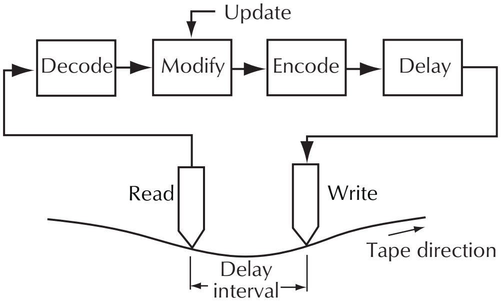


Figure 28-45. Block diagram of the audio system of a typical digital recorder.

Some DASH recorders offered analog to digital conversion as extra cost options. Generally, digital data was fed through the Digital In port. This data was immediately routed to the output for monitoring via the Input Select selector. The data were also fed to a Crossfader that smoothly switches between the input source and tape playback for punch-ins. The data to be recorded was first spread out by the Interleave circuit to minimize the impact of a burst error. A powerful Reed-Solomon error correction encode process was applied by the RSC Coder. The data was then delayed by a variable amount to account for fixed and variable timing errors. A 4/6 Modulator adds bits to the data to create easily recognized data patterns that have an optimized bandwidth for recording on the tape. The patterns were then fed through the Write Amplifier to the write head.

Playback from the tape began at either of two read heads, one located before the write head for sync/overdub operations and one located after the write head for confidence monitoring. The selected data was fed to the Read Amplifier where the analog-looking pulses from the tape were converted into digital pulses by a differentiator and level detector. The patterns of digital pulses passed through the 4/6 Demodulator to strip off the extra bits added by the modulator during the write process. The Timebase Corrector removed any timing variations due to flutter in the tape transport, restoring the exact sampling rate. The RSC Decoder and Deinterleave used the Reed-Solomon data to correct any correctable errors and put the data back into the proper order. Any uncorrectable errors were concealed by the Interpolater, which made a best guess attempt to hide errors. If the errors are too large to hide, the output mutes rather than passes faulty data.

Other functions include master timing circuits, tape motion servos, extensive logic, and display functions for metering and control.

28.5.2 Helical Scan Digital Tape Transports

The primary limitations of longitudinal magnetic tape with a large number of narrow tracks were tape guiding and crosstalk and of course expense. Helical scan techniques greatly reduce these problems by using heads with alternating azimuth angles on the record/play heads to reduce crosstalk. Track-to-track spacing can be virtually overlapping, with dynamic tracking of the flying heads to eliminate any errors, Fig. 28-46.

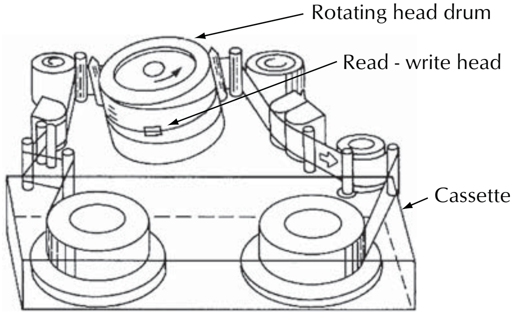


Figure 28-46. Helical scan transport layout.

The limiting factor for helical scan tape is the throughput on the single digital track that is being recorded. For example, a recorder capable of 8 channels of 16 bit or 24 bit audio requires bandwidths of approximately 8 MHz or 12 MHz, respectively. The most economical approach is to adapt a consumer format to fit this requirement. For example, the popular ADAT series manufactured by Alesis and others were based on the S-VHS format that used ½ inch tape. Similarly, the Tascam DTRS (Digital Tape Recording System) series utilizes the technology developed for 8 mm handheld video recorders. Both of these products offer 8 channels in an inexpensive package. Multiple machines can be locked together to provide up to 128 tracks of audio at about one-tenth the price of the equivalent DASH recorder.

The basic helical scan transport consists of a rapidly rotating head drum and a capstan to control the forward speed of the tape. A spooling mechanism engages the reel hubs in the cassette to provide proper winding of the tape in all modes. Auxiliary functions include auto loading mechanisms to load and eject the cassette and auto threading mechanisms to extract the tape from the cassette.

For most applications, the tape is wrapped around the head drum to cover slightly more than half the circumference of the drum. Heads are mounted in pairs 180° apart on the drum, protruding slightly from the face of the drum. The specific tape format determines the diameter of the head drum. The drum spins many revolutions per second to provide the high linear scanning speed required for the digital data stream. For example, in 16 bit mode the ADAT drum is 2.44 inches in diameter and spins at 50 rev/s to yield a linear scanning speed of 192 in/s. In comparison, the forward speed of the tape is only 3.9 in/s, about one-fifieth of the scanning speed.

The scanning drum is tilted slightly with respect to the path of the tape, causing the spinning head to scan the tape in diagonal stripes. For the ADAT example, the angle is 7.5°, yielding a diagonal track length of slightly less than 4 inches. In comparison, the R-DAT system uses a 1.18 inch (30 mm) diameter drum inclined 6.5° spinning at 33.3 rev/s to give a scanning speed of about 120 in/s.

28.5.3 Rotary Digital Audio Tape

The R-DAT format shares technology with the 8 mm camcorder VCRs. By optimizing a miniaturized helical scan system using 4 mm tape for direct digital recording of stereo audio, a very compact digital recorder, Fig. 28-47, has been made possible. New heads and metal particle tapes have been utilized to produce a long-playing cassette tape system with quality equal to the compact digital disk.

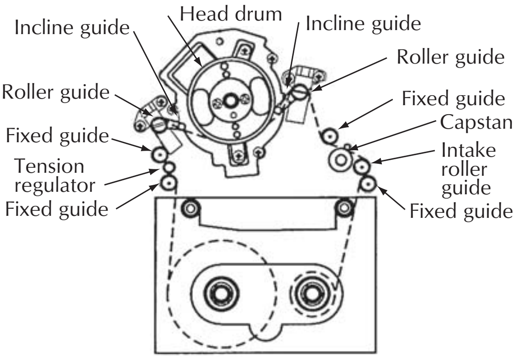


Figure 28-47. R-DAT tape transport mechanism.

The R-DAT format operates at two tape speeds, 8.15 mm/s (0.32 in/s) for recording and 12.23 mm/s (0.48 in/s) for widetrack playback of prerecorded tapes. In spite of the very slow tape speed, very high data rates are made possible with a rotating head drum speed of 2000 rev/min and flying head velocity of 3 m/s. The resulting slant tracks are 23.5 mm (0.93 in) long and inclined at an angle of approximately 6.5° from horizontal.

The data is digitized to 16 bit resolution and recorded with double-encoded Reed-Solomon error correcting coding with interleaving between not only channels 1 and 2 but also adjacent scans of the flying heads, Figs 28-48. A 60 m (65.6 ft) tape holds 2200 Mbytes of information capable of encoding 2 hours of stereo music. Search for a desired program can be conducted at 60 times normal speed; rewind and fast forward without search is up to 180 times normal speed, allowing full rewinding in approximately 40 s.

The digital storage capacity of the audio R-DAT format has been greatly enhanced by newer technology to serve as a backup medium for computer hard disks. The fourth generation of the Digital Data Storage format (DDS4) jointly developed by Sony and Hewlett-Packard from the original R-DAT format boasts a capacity of 20 Gbytes per 150 meter tape. The drum spins at 11,480 rev/min to achieve a data throughput of 2.87 MB/s before data compression and up to 7.62 MB/s after compression.

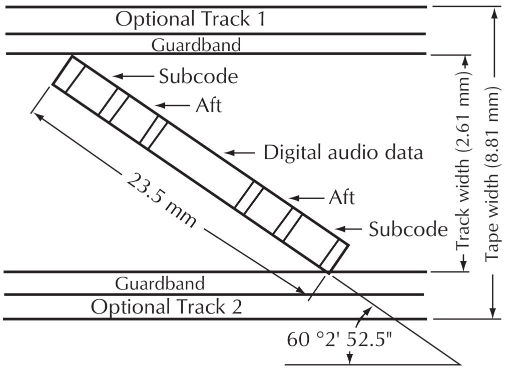


Figure 28-48. R-DAT tape format.

28.5.4 Packing Density Maximization with Rotary Head Recorders

The professional analog multitrack formats are very inefficient in the use of recording tape. Nearly half of the tape width is devoted to guardbands between tracks. These guardbands are required to minimize crosstalk between channels due to fringing and crosstalk within the heads. These problems are overcome in rotary head recorders that do not require guardbands.

Several aspects of the rotary head system contribute to the elimination of guardbands, including servo positioning of the tape, azimuth shifting on alternate scans, and the elimination of low-frequency components in the recorded signal.

The servo positioning of the helical scan tape during playback is analogous to a conventional longitudinal recorder with self-aligning guides to correct for any guiding errors. Control signals recorded along the edge of the tape are used to synchronize the motion of the tape past the rotating drum to the spinning of the drum. This synchronization adjusts the position of the recorded tracks on the tape to exactly coincide with the path of the flying head. The active servo control of the tape motion duplicates any disturbances that may have occurred during recording to maintain correlation between the flying head path and the track, permitting tracks to be recorded abutting each other.

This technique can be carried one step further if the spinning head is augmented with a rapidly responding positioning actuator. Fig. 28-49 shows a scanning head mounted on a piezoelectric positioner called a *bimorph*. If a voltage is applied to the bimorph, the head mount deflects and moves the head.

Since the bimorph can respond much faster than the servo system, tracking errors can be continuously corrected throughout the helical scan of the tape. This

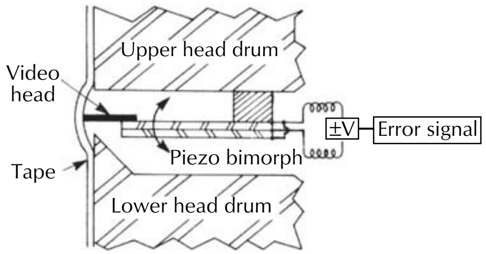


Figure 28-49. A dynamic head positioner using a bimorph.

wider bandwidth, however, requires a method of actually sensing any errors when the head begins to slip off the center of the slant track. One of several techniques for this purpose is utilized in the 8 mm helical format. As shown in Fig. 28-50, low-frequency tracking signals are added to the high-frequency data. Four different frequencies are recorded on four sequential passes, with the frequencies chosen so that the difference between adjacent frequencies is either 16.5 kHz or 46.2 kHz.

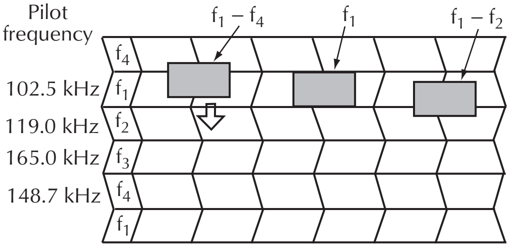


Figure 28-50. Pilot signal tracking servo showing pilot frequencies.

If the video head on track f_1 plays back the signal accurately, only the signal f_1 (102.5 kHz) is reproduced. If the video head deviates toward the f_4 track, it will pick up both signals f_1 and f_4 (148.7 kHz). The difference between these two signals, $f_4 - f_1$, will give an error signal Δf of 46.2 kHz, causing the video head to be moved back toward the f_1 track immediately. If the video head shifts to the f_2 track, another error signal $f_2 - f_1$ of 16.5 kHz will be produced, and the video head will be moved back toward the proper f_1 track. Thus, the video head can be made to accurately follow a previously recorded track.

The close packing of adjacent tracks would cause serious fringing problems if the data signals on the tracks contained any long-wavelength information. To avoid any such problems, the audio signal is encoded using either digital or FM techniques to shift the frequency content upward and eliminate all low frequencies.

Isolation of the short-wavelength encoded signals between adjacent odd and even tracks is further improved by offsetting the azimuth tilt of the heads during recording and playback as shown in Fig. 28-51. The resulting azimuth error for any signal leaking from the adjacent tracks will partially attenuate any crosstalk.

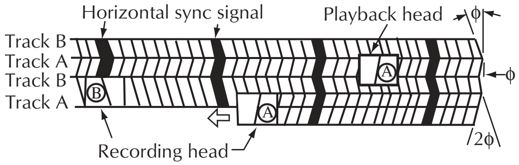


Figure 28-51. Differential azimuth recording technique.

A helical scan recorder must have additional circuitry to assemble the digital data from several tracks into a serial stream that is recorded as blocks of data by the scanning head. The data must usually be replaced as an entire block, necessitating a complete rewrite of all channels if any channel is changing.

All of the digital circuitry of a helical recorder can be squeezed into just a few custom integrated circuits. The newer generations of ADAT machines, for example, adopted digital servos for controlling the transport so that all of the motor servos could be consolidated into a single chip, eliminating the need for any analog servo adjustment potentiometers. The digital signal chain is also highly integrated, resulting in an amazingly uncomplicated main circuit board with just a few ICs for the entire machine.

28.5.5 Heads for Digital Tape Recorders and Hard Disk Drives

The packing density of the data on hard disks in 1990 was around 100 mb/in^2. At the time of this publication, it is over 200 Gbits/in^2**.** Fig. 28-52 shows a thin film digital tape head.

As the areal density of the data on tapes and disks increases, each bit must shrink in size. The smaller bits contain less magnetic energy and generate smaller electrical pulses in the coil of a read head. The resulting loss in SNR eventually imposes a useful lower limit on the size of the bits.

This limit has been pushed back by read head technology called *giant magnetoresistive* (GMR) or *spin valve* heads. (The term *giant* differentiates these very high-output heads with giant output signals from earlier low-output magnetoresistive heads.) The GMR head is fabricated by vacuum deposition, creating a sandwich of metals that changes resistance when excited by a

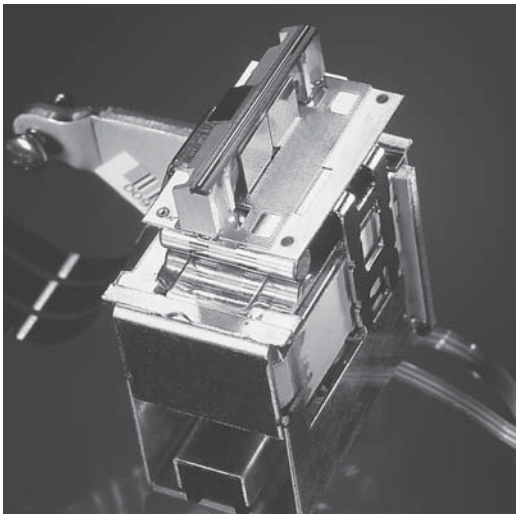


Figure 28-52. Thin film digital tape head.

magnetic field. The rear layer of the sandwich shown in Fig. 28-53 has a fixed or *pinned* magnetic field that serves as a reference. The *filler* of the sandwich is a magnetoresistive (MR) material chosen for a large change in resistance per change in magnetic flux. The front outer layer is a magnetic probe that actually samples the magnetic flux of the bits on the disk. As the magnetic polarity of the data bits reverses, the angle of the magnetic field in the outer layer *spins* back and forth. Part of this field bridges through the center layer to the pinned rear layer, causing the resistance of the MR layer to change. The resulting output signal has a much better SNR than an equivalent read head.

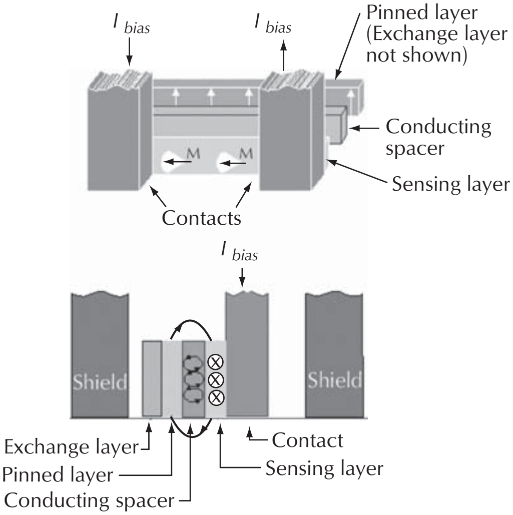


Figure 28-53. Giant magnetoresistive head.

Since the GMR effect does not work in reverse to generate a varying magnetic field when driven by an electrical signal, we still need a coiled conductor for writing the data onto the disk. The solution is a composite head that has both a GMR read element and a coil for writing. The entire head, including the GMR read element and the coil for writing, can be fabricated together using thin film techniques. A single thin film wafer may contain up to 20,000 heads.

Most digital recording schemes drive the record head hard enough to saturate the medium in one polarity or the other. If the head is tracking exactly over any prior data, the old data will be completely overwritten. Unfortunately, the tolerances of the head tracking system may cause slight alignment errors that leave a bit of the old signal unerased.

One method to remove the residue is to use a *straddle erase* technique that resembles the outriggers on a Hawaiian canoe. Two thin erase cores straddle the desired track and trim off any of the prior signal that wasn't covered by the new recording.

A newer technique is to *write wide and read narrow*. Just as we discussed for analog recording, we can write a track that is wider than the read core. The extra width of the recorded track allows for a small tracking error. This technique is easily implemented with GMR heads since these heads have separate read and write elements. The read element is fabricated slightly narrower than the write element to create the desired overlap.

28.5.6 Magnetic Disks

A 2500 foot roll of 2 inch recording tape has enough surface to carpet a large living room. A 60 minute DTRS tape would only cover half of a couch. In comparison, a multigigabyte hard disk in a digital audio workstation uses a few 3½ inch (89 mm) diameter magnetic disks with a working area about the size of your footprint. Although the basic technology of all of these products is similar, the precision required in their manufacturing increases rapidly as the size shrinks and the density increases.

28.5.6.1 Floppy Disks

Floppy disks were close cousins to magnetic tape. Although the disks were cut from large rolls much like the jumbo rolls from which magnetic tape is slit, the coatings parameters were very different.

The diskette, which we now call the floppy disk, was developed around 1970 as a read-only device. The contents of the prerecorded diskette were loaded into a computer or storage system to furnish startup or diagnostic information, much like a boot ROM in today's computers. Over the next ten years the product evolved from an 8 inch diameter single-sided read-only device to a 3.5 inch double-sided read/write device with twenty times the capacity of the original diskette.

Although the original diskette operated on only one side of the disk, the media had magnetic coatings on both sides to promote flatness. The symmetric construction of the disk was eventually exploited to double the data capacity by recording on both sides.

By the end of their evolution, floppy disks utilized a very thin coating of cobalt doped gamma ferric oxide. The coating thickness was about one-fifteenth the thickness found on our analog mastering tapes, and the particle coercivity was about twice as high.

One important difference in floppy disk media is that the magnetic particles for a spinning disk must not be oriented in a single direction as on our audio tapes. Recording characteristics degrade several dB when an oriented tape is operated crosswise to the intended direction. A linearly oriented disk would therefore see large peaks and troughs in output at twice the rotational speed. To avoid these fluctuations, the floppy disk coating process is optimized to either disperse the magnetic particles in a random orientation or orient the particles circularly.

The floppy disk operated with the magnetic head in direct contact with the magnetic media, just as in a tape recorder. As a result, the floppy disk system was subject to head wear and head clogging due to dirt and debris.

28.5.6.2 Rigid Disks

Tape and floppy disk recorders utilized contact recording with the head touching the recording medium. This continuous sliding contact produced wear that limited the life of the heads and medium. The flying head of a hard disk drive eliminates this contact, greatly extending the life of the head and disk.

High digital densities require very low flying heights on the order of 0.4 µinch (10 nm) to avoid excessive spacing loss. Any surface irregularity that sticks up more than the flying height will impact the flying head. To make matters worse, even smaller defects can upset the aerodynamic flow of the air around the head enough to cause instabilities that can make the head crash into the disk surface. To avoid these problems, the disk substrate and the magnetic coating must both be extremely smooth.

Aluminum/magnesium alloys and glass are the preferred substrate materials, with glass rapidly gaining popularity as disk sizes decrease. Plastic disks, some with servo patterns pressed into the surface during the molding process, are also entering the market.

Aluminum disk substrates are cut from special aluminum sheet that is optimized for flatness and surface smoothness. The disks are polished and then plated with an undercoat of nickel phosphorus (NiP).

Glass and glass/ceramic substrates are rapidly displacing aluminum disks. Glass offers a very smooth surface and a higher stiffness than aluminum. The benefits are a lower flying height with fewer surface defects and a disk that is more robust.

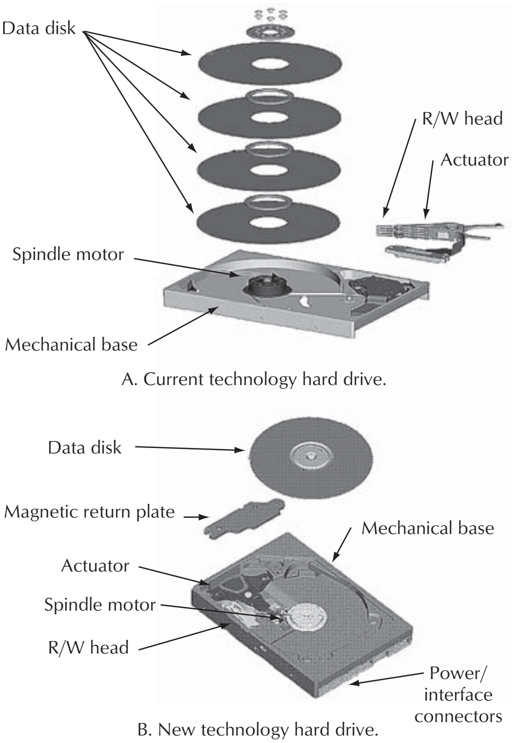
The aluminum or glass disk is coated with multiple layers that include foundation layers, the active magnetic surface and protective overcoats. Although earlier disks were spin coated with a slurry resembling the coating for magnetic tape, modern disks are prepared by plating and ion bombardment. Hard diamond-like overcoats and surface lubricants protect the magnetic layer from accidental contact with the head.

28.5.7 Hard Disk Drives

Rotating disk drives offer very rapid random access to a huge array of data, Fig. 28-54. This yields two very important benefits. First, the rewind, fast forward, and autolocate functions of a tape recorder become nearly instantaneous. This speeds up operation, especially during editing sessions.

A second and much more important benefit is the ability to rearrange the output data. Assuming a fast host computer with versatile digital audio workstation (DAW) software, the user can construct a song from a multitude of track segments almost as if he or she cut each track of a reel of multitrack tape into a thin ribbon and then chopped and spliced the individual ribbons back together to arrange the song. This incredible versatility has fueled the rapid replacement of analog audio tape recorders in recording studios. Even when an analog recorder is employed for the initial capture of the music, the analog tracks will probably be digitized and loaded into a DAW for editing and mixing.

To demonstrate how data are stored on a spinning disk, consider the inner workings of a representative single-platter drive. Although this unit is only a single-platter, 15 Gb entry-level drive under $100, this drive's areal density of 22.5 Gb/in^2 led the industry when the drive was introduced in early 2001. We will look at the major subsystems to rotate the disk, position



A. Current technology hard drive.

B. New technology hard drive.

Figure 28-54. Interior of a hard disk drive.

the read/write/erase head assembly, and process the data to and from the disk.

The spinning disk is an aluminum or glass disk covered with a magnetic layer. Since smaller disks are flatter and more rigid, the trend has been downward in disk size from 14 inch diameter disks in 1960 to disks ranging from 3.5 inch down to 1.8 inch diameter today. Smaller disks can spin faster and rotation rates have risen from about 3000 rpm for the 14 inch disks to targeted speeds of 22,000 rpm for high-performance small disks.

These higher rates and tighter tolerances are exceeding the capabilities of the ball bearings that support the spinning disk, requiring new types of bearings. Fluid dynamic bearings replace the rolling balls in a ball bearing with a film of oil that is less than one-tenth the thickness of a human hair. In addition to providing tighter tolerance, the fluid dynamic bearing is quieter, longer lasting, and more rugged.

The spindle assembly includes an integral motor for spinning the disk. The power required from the spindle motor due to aerodynamic drag of the spinning disks is

$$P \propto n \times \omega^{2.8} \times r^{4.6} \tag{28-11}$$

where,

n is the number of platters,

ω is the angular velocity of rotation,

r is the disk radius.

Additional power is required to overcome the friction and viscous losses of the bearings.

The magnetic data on the disk is accessed by magnetic heads flying over the surface of the disk. The very close spacing between the head and disk is maintained by a cushion of air generated by the aerodynamic design of the head. For our example drive, the head-to-disk spacing is 0.6 μin (15 nm) or 1⁄40 the wavelength of orange light.

Since the head cannot fly when the disk rotation stops, provisions must be included to transition from flying to nonflying status. Some drives land the heads on a dedicated portion of the disk periphery appropriately known as the landing zone. Other drives move the head to an extreme position to engage a parking ramp that holds the head away from the disk. The parking ramp also provides protection from shock and vibration incurred during shipping or handling of the computer.

Some of today's disk surfaces are so smooth that the head will literally stick to the surface after landing. To overcome this *stiction*, the surface of the disk may be textured with microscopic bumps. The bumps may require an increase in the flying height, thereby reducing the maximum storage density of the disk.

To avoid these problems, the example drive uses a parking ramp at the center of the disk. The resulting ability to use an untextured disk surface is a major reason for this drive's very high packing density.

The flying head is mounted on metallic spring matrix called a *gimbal* that allows the head to assume the proper flying attitude parallel to the disk surface. The gimbal is at the end of a long support arm called a *flexure* that cantilevers the head above the disk surface. The flexure lightly presses the head onto the disk surface to overcome the aerodynamic lift generated by the head.

The flexure is attached to an actuator that moves the head to the appropriate track of magnetic data on the disk surface using either linear or rotary motion. The linear actuator is very similar to the voice coil and magnet of a loudspeaker. A current in the coil produces a magnetic field that interacts with the field of the permanent magnet to create a linear force along the axis of the coil. A sled assembly with ball bearing wheels maintains alignment of the coil as the head moves in a straight line along a radius of the disk.

A rotary actuator is a pivoting device that moves the head in an arc across the surface of the disk. The arc causes the head to depart from absolute tangency to the disk, but the error is relatively small since only 30% of the disk's radius contains data. The actuating force is once again generated by a coil and magnet, but in this case the components are curved to match the pivoting motion. The support bearings and structure of the rotary actuator are simpler and less expensive than the linear actuator's sled.

The actuators are part of a closed loop control system that moves the head to the appropriate radius on the disk. The desired address comes from a translator that converts a data address into a physical radius. The feedback to the control system is the actual data that is being read from the disk. Older systems used one surface of one disk in the stack of disks for nothing but positioning information. These *dedicated servo* drives contained prerecorded positioning information defining the radius and the angle of rotation. The actuator would move to match the desired radius to the data being read from the servo platter.

Newer drives save the expense of a servo platter by using information written on the normal data surfaces. These *embedded servo* schemes have short blocks of address information scattered around each circular track at regular intervals. These systems can also sense changes in the readout pattern when the head begins to move off the centerline of the data track.

Embedded servos have allowed designers to greatly increase the track density on the surface of the disk. Any small changes due to temperature and wear are actively sensed at the exact point where the data is being written and read, not at a remote location on another disk. Our example drive uses an embedded servo to pack 40,000 tracks per inch of radius.

The current trend is to utilize digital signal processor (DSP) chips for the tracking servo and spindle motor control. Other electronics tasks include a buffered digital data interface with the host computer, error detection and correction and encode/decode of the data to optimize the read/write process.

All of the mechanisms and servos in a disk drive would be useless if any dirt gets into the system. When a head is flying at a spacing of less than 1 µinch, even a particle of cigarette smoke can cause a catastrophic collision that might destroy the head and/or disk. To avoid contamination problems, the entire head and disk assembly are enclosed in a clean environment. Any air entering the sealed head/disk assembly for cooling or atmospheric pressure balancing passes through a filter that traps all dirt.

28.5.8 Hard Disk Electronics

The hard disk drive also features very dense circuit packaging. A typical drive has less than 20 in^2 of circuit board with just a few highly integrated chips. The block diagram of Fig. 28-55 shows the basic functions that are squeezed into this small space.

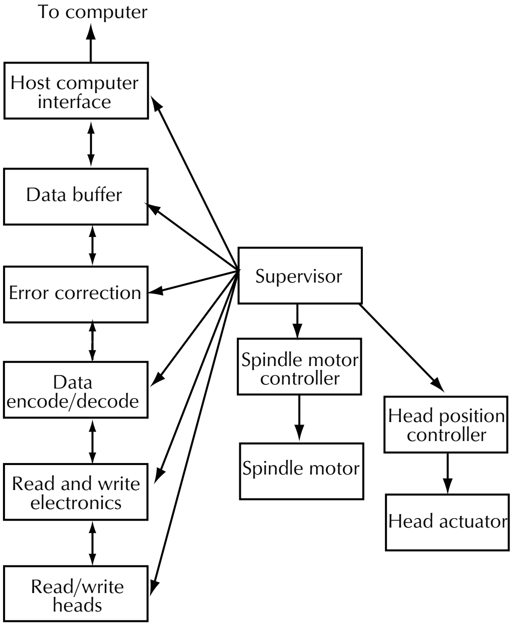


Figure 28-55. Block diagram of a hard disk drive.

The spindle motor controller provides the servo loop that turns the disk at a constant speed. The controller also provides the acceleration and deceleration profiles during startup and shutdown.

The actuator servo controls the linear or rotary voice coil motor that positions the head actuator at the proper radius of the disk. This servo must provide rapid seeks to the desired data and track any eccentricities or other disturbances that might cause a tracking error. The current trend is to program custom DSP chips to serve as digital servos for both the actuator servo and spindle motor controllers.

The data path circuitry provides many of the interleaving, error correction, and modulation code functions described in conjunction with the digital tape recorder above. In addition, the interface provides data formatting and buffering to handshake with the host computer system.

The hard disk data path is much faster than the DASH tape recorder. Read/write circuits for hard drives are pushing frequencies of a gigahertz, yielding data rates of 100 megabytes/s. At 3 bytes per 24-bit audio word and a sample rate of 96 ksamples/s, this represents about 250 audio channels. This number is best case, and the throughput drops drastically if the drive must read and write simultaneously while seeking various tracks of data.

The trend is toward higher levels of integration in disk drives, with more of the very high-speed circuitry moving closer to the read/write head to avoid delays and waveform distortions due to wire lengths and inductances. (Electricity travels about 1 foot in a nanosecond, and 1 ns is the period of one cycle of a gigahertz signal.)

The adoption of fluid dynamic bearings permits higher disk rotation speeds that reduce the latency time for a desired block of data to rotate to the head's location. The average latency is half the rotation period for the disk. For a 15,000 rev/min (250 rev/s) disk drive, the average latency is 2 ms.

28.5.9 Formatting Media

Digital media typically require one or two stages of preparatory recordings of control information before user data can be recorded. *Low-level formatting* involves basic housekeeping tasks that allow the drive to properly and accurately move the media and heads to the correct physical locations. In addition, *high-level formatting* defines the nature of the digital data blocks regarding sector and block lengths. Formatting also checks for media defects, marking bad sectors and relocating data to good sectors.

Formatting may also include writing control tracks with synchronization and address information. To illustrate why, consider a helical scan digital audio tape recorder. We can start with a blank tape and begin a recording. The machine records the helical stripes of data with embedded address information and a control track along the edge of the tape to facilitate synchronizing the linear tape speed with the rotations of the helical drum during subsequent playbacks. If we stop the recording, we also stop the recording of all of the address and synchronizing data.

If we wish to restart the recording by punching in at our previous exit point, we must seamlessly append address and synchronization data to the ends of the previously recorded tracks. But what happens if the recorder is running at a slightly different speed, perhaps due to the recorder warming up, when we punch in? Whenever we play back the tape, the recorder must abruptly change the tape speed at the punch-in point.

A better technique is to prerecord or *format* the entire tape with address and synchronizing information. This will allow us to locate any address on the tape in a continuous manner, and the tape speed will be constant throughout the tape.

Formatting a tape or disk can be a very time-consuming task. Tapes typically must run through the machine at normal speed; hence a 30 minute tape would require 30 minutes for formatting. Preformatted tapes with prerecorded address and synchronization tracks are now available from the tape manufacturers for some of the digital audio formats. In addition to saving time, the preformatted tapes also reduce wear on the recorder.

Hard disk formatting may vary from minutes to many hours, depending on the operation. The lengthiest operation is *repacking* all of the data on a disk. After a file has been changed a number of times, the physical file may be many sections scattered widely across the surfaces of the disks, leaving unusable islands of updated and deleted data. The repacking operation relocates and reassembles the files as contiguous data, freeing up the wasted space. Repacking also checks the disk surface for defects. If a sector is contaminated with errors, the drive may try several read operations to recover the data. Some drives will also move the head off-track slightly to recover poorly written tracks. Bad sectors are marked so that they will not be reused.

Although audio tape recordings do not contain any addressing and synchronizing information embedded within the audio recording, some applications require adding a track of SMPTE/EBU timecode for synchronization or editing. For live production work, the timecode track will be recorded on all of the audio and video machines to allow later synchronization of multiple machines.

Timecode is also used during the editing process to identify segments that are to be assembled onto a master reel. Timecode is first prerecorded or *striped* onto the master reel to allow the editing computer to precisely locate the destination addresses of all of the edits. The computer then locates the appropriate segments in the timecodes on the source reels and copies the audio and/or video onto the designated timecode section of the master reel.

28.5.10 Long-Term Storage

A common question is "How do I store my digital data when I finish a project?" Many users have decided that hard disk drives are cheap enough to just store the hard drive as the archival copy. This strategy is fraught with problems that could come back to bite the user. Disk drives have several failure mechanisms that can render the data unrecoverable over time.

Some systems park the head on the surface of the disk. Over time, the lubricant that is embedded in the disk's coating can migrate to the surface and "glue" the head to the disk. This problem is avoided with drives that have parking ramps to hold the heads off the surface of the disk when the disk drive isn't running.

Another problem is the spindle bearings. If the drive is stored for extended periods, the lubricant may degrade or migrate away from the critical bearing surfaces. This can lead to bearing failure when the drive is restarted.

The manufacturer rates a typical drive for about three years of useful life. There is no separate specification regarding storage life. Expecting a long storage life is an act of sheer faith.

The advantage of a tape or optical disk backup of the digital data is that the media and the mechanism are two separate items. The drive mechanism can be maintained and serviced without involving the media. The problem then becomes finding a working sample of the appropriate drive, or finding parts and a trained technician to fix a nonworking sample. Several digital tape formats have already reached the point at which finding a working tape deck to play the tapes is difficult or impossible. This problem will only get worse in the future.

If the data is valuable, the user should map out a backup strategy that will assure accessibility. This may require occasional copying of the digital information to newer formats. If nothing else, the user should have a schedule to verify every year or two that the original data can still be accessed without any degradation.

28.5.11 Data Interchange

Standards for compatibility of digitally recorded tapes have become much more difficult to achieve because of the wide range of choices open to the digital audio designer. The common problems of mechanical compatibility of tape speed and track format are still present, plus the sampling rate, data format, timing, and error-handling methods must also be compatible.

The rapid evolution of digital audio technology in these areas, which has already rendered several generations of digital audio recorders obsolete, has blunted any attempts at standardization at the media level. The point of data compatibility has moved up to the electronic interface between systems. At this level we find widely used standard protocols such as AES/EBU, SPDIF, and ADAT light pipe. Additional work, such as AES 31, to standardize file transfer protocols between hard disk systems will provide for the electronic transport of audio files throughout a facility via local area networks, and throughout the world via the Internet.

28.6 Tape Recorder Transport, Maintenance, and Testing

Maintenance begins with inspection and cleaning. Before starting the cleaning procedure, note the location and type of dirt and debris that has accumulated due to prior use. Excessive debris indicates that your recording tape is being slowly destroyed by the tape transport.

A deposit of very fine, silky threads indicates that the polyester base film of the tape is being scraped off by a sharp edge on a guide flange. Examine all edge guides for grooves cut into the flanges by the tape. Either reposition the guide to place an unworn surface in contact with the tape or install a new guide if the groove is severe.

Deposits of brown or black dust near the guides indicate that the edges of the tape are being scraped or deformed enough to break small chunks of coating from the edge of the tape. Check the tape tension and the height of the guides and reel hubs.

Any caked-on deposits on the surface of the guides or heads are very serious. Inspect the surface of the tape for scratch marks. If the tape surface is being scratched, continued use will destroy the tape. Correct the cause of the scratches before continuing.

Several types of cleaners are available for cleaning tape machines. Older head cleaners usually contained Xylol, a strong solvent, to aggressively dissolve tape residue. Milder isopropyl alcohol is a more popular solvent today, but avoid rubbing alcohol containing 30% water in favor of the 99% pure variety for topical antimicrobial use.

Use a soft swab moistened with cleaner to scrub the contact surfaces of the heads, guides, and capstans. Avoid drenching the swab. If the swab is too wet, solvent may run down the capstan shaft into the top bearing, washing away the bearing's lubrication. Cotton swabs are suitable for most analog tape recorders but not for the delicate heads on a helical scan recorder. Use special lint-free swabs with more pliable sticks for cleaning rotary head machines.

When cleaning the head, always rub the swab in the direction of tape motion, never across the head sideways. Sideways scrubbing may peel away the edge laminations of the cores. Avoid scraping the face of the head with the stick or core of the swab. Allow adequate time, typically 30 s, for the solvent to evaporate before rethreading the tape. You don't want the leftover solvent dissolving your recording tape.

Xylene head cleaning solvents will attack some plastics including the lenses of optical sensors. Aggressive solvents may either partially dissolve or create a hard, glazed surface on some rubber rollers. If you notice a lot of reside on your swab or rag after wiping a seemingly clean roller, you are probably dissolving the roller, not cleaning it! Use general- purpose cleaners for the plastic components and rubber cleaners for the rubber rollers.

The tape must also be kept completely free of dirt. Keep the surface of the transport clean to avoid dirt being picked up during high-speed spooling. Always return the tape to its storage carton between uses. Do not stick your fingers through the windage cutouts in the flanges of the reel and touch the edges of the tape pack when handling the reel. (Skin debris from fingers is a source of tape dropouts!) In addition:

1. Avoid eating greasy foods while handling tapes.
2. Contamination due to finger oils and debris can be avoided during editing sessions by wearing lint-free editing gloves, which are available at most camera supply stores.
3. Keep cigarette ashes and other powdery materials far away from the tape.

The cooling system of the tape recorder should be cleaned periodically. Clean all air filters and cooling passageways and remove any dust buildup with a vacuum cleaner. Verify that all inlet or exhaust ports on the bottom of the machine are not obstructed by carpeting or dust and that adequate clearance for free airflow exists at the rear of the machine.

Following cleaning, diagnostic servicing should begin with verification that the tape guiding and tension at the heads is adequate to maintain good tape-to-head contact. Set aside one reel of tape, known as a *shop tape* because it typically comes from the maintenance shop, for testing. Run this tape in all modes while observing the tape at the heads and the guides. The tape should not run hard against either guide flange and there should never be any edge distortion. If edge distortion is noted, check for a bent guide or tension sensor arm. These components can be easily bumped out of alignment by a full reel of tape during loading or unloading.

On many machines a tape tension gauge of the type shown in Fig. 28-56 can be inserted in the tape path near the heads to measure the tension. For other machines that are too crowded in the head area, either the head assembly must be removed or a test location away from the heads must be used. Measure the tension at both the beginning and the end of the reel.

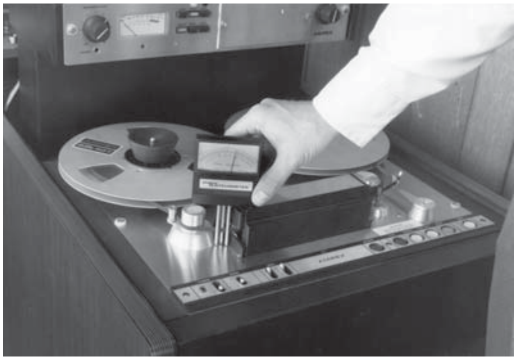


Figure 28-56. Tension measurement. Courtesy Ampex Corp.

Note that the stiffness of a piece of tape varies with the width, base film thickness, and type of tape. The tension gauge must be adjusted before use to read correctly for the specific tape sample being used on the transport. A calibration weight is included with the gauge for this purpose.

The following tape tension values indicate the range of tensions commonly encountered on studio recorders.

¼ inch	3–4 oz
½ inch	4–8 oz
1 inch	6–12 oz
2 inch	10–24 oz

The nominal value for a given model of recorder will be found in the maintenance manual for the machine.

Some manufacturers specify tension measurements with a spring scale and a cord that is wrapped around a tape hub. Follow the recommended procedure.

Verify that the mechanical brakes or dynamic braking logic is stopping the tape smoothly from all modes and speeds without excessive force. A sticky brake solenoid or dirty brake band can quickly ruin your precious tape.

28.6.1 Speed

Absolute tape speed is extremely difficult to measure, even under the controlled conditions of a laboratory. One method available to maintenance personnel is to measure the frequency reproduced from a commercially available speed reference tape. The frequency read on the frequency counter must be corrected for any difference between the tape tension on the playback machine and the tension value used by the manufacturer of the tape during the recording process. A correction table is furnished with the tape for this purpose.

A more common speed test is to check speed uniformity from beginning to end of a reel of a tape. The following procedure outlines the general technique:

1. Using an oscillator that has been operating long enough to reach stable conditions, record a reference tone in the range of 1–5 kHz at the head end of the tape.
2. Flip the reels so that the head end becomes the tail.
3. Using the console monitoring provisions of the console, mix the reproduced tone with the oscillator tone, listening for any major pitch differences. (If a significant error is detected, flip the reels again to verify that the oscillator has not shifted frequency.)

A more accurate version of this test is to use a frequency counter to measure the frequency at both ends of the reel. The speed error in percent is then calculated as

$$\%\text{ speed error} = 2\left(\frac{head - tail}{head + tail}\right) \times 100\% . \qquad (28\text{-}12)$$

A speed error of 6% will yield a pitch change of one half-tone step. Typical recorder specifications are in the range of 0.1–0.5%. Machines with constant tape tension will generally have the least error.

Possible causes of speed error include excessive tension variations from beginning to end of the reel, tape slippage due to a worn capstan surface or pinch roller, inadequate pinch roller pressure, and unstable capstan speed.

Assuming that tape tension has already been determined to be correct on both sides of the capstan, the next test is to check pinch roller pressure. First, inspect the pinch roller for glazing of the roller surface or excessive wear. Fig. 28-57 shows roller wear patterns that may reduce the traction between the tape and capstan.

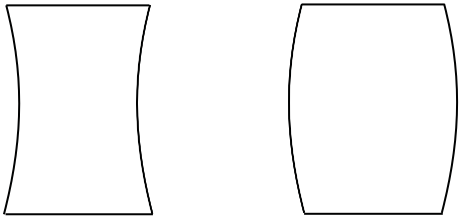


Figure 28-57. Pinch roller wear patterns.

Next, a spring scale is coupled to the top (and the bottom, if possible) of the pinch roller yoke or arm, as shown in Fig. 28-58. The scale is pulled at right angles to the support arm with just enough force to disengage the roller from the capstan. The force reading at disengagement should be compared with the recorder manufacturer's recommended value.

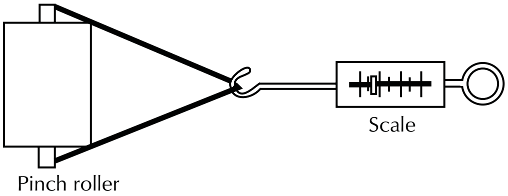


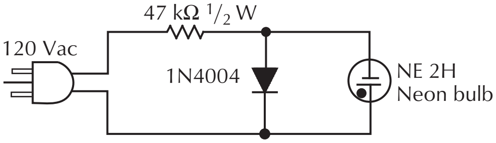
Figure 28-58. Pinch roller force measurements.

For some transports the pinch roller force is set as a fixed number of turns of a nut or screw. For this case the roller linkage is first tightened to bring the roller into light contact with the capstan, and then the recommended clamping force is applied by tightening the adjustment by the specified number of additional turns.

The surface of the capstan may become so highly polished by the abrasive action of the tape that slippage will persist for the correct values of tension and pinch roller pressure. In this case the capstan must be resurfaced by plating or sandblasting or both to restore the required traction.

In very rare cases the capstan motor may actually slow down due to excessive loading caused by bad motor bearings or high tension. Bushing bearings, which are used on many direct-drive ac synchronous capstan motors and some capstan pinch rollers are an especially noteworthy problem. Periodic lubrication of these components is essential to maintain low-friction operation. Although these components may appear to spin freely when turned by hand in an unloaded state, the friction can rise dramatically when the engagement solenoid exerts several pounds of side load on the bearings. The resulting drag and wear due to dry bearings may produce substantial speed errors. One small drop of oil can make all the difference in the world. To avoid problems, follow the manufacturer's recommended lubrication schedule.

A simple strobe light, as shown in Fig. 28-59, can be used to check the running speed of the flywheel or fan on the shaft of the synchronous capstan motors. Package the components inside a discarded plastic pen housing with the tip of the bulb protruding. Hold the light close enough to the rotating device to observe a reflection. The reflected pattern must remain stationary under all conditions of tape pack and speed. Induction motors, which do not run at synchronous speed, will always yield a moving pattern.



Note: Package the components inside a discarded plastic pen housing with the tip of the bulb protruding.

Figure 28-59. Strobe light for speed testing.

Crystal-referenced servos may falsely appear to vary in speed when tested with a strobe light if the frequency of the ac mains driving the strobe varies. An oscilloscope and frequency counter are required to properly verify correct servo operation.

28.6.2 Flutter

Speed drift represents only the very lowest frequency components of the spectrum of speed errors. Measurement of the higher-frequency flutter components requires a specialized frequency demodulating instrument called a *flutter meter*. As seen in Fig. 28-60, the flutter meter may resemble the phase-lock servo of Fig. 28-4. The reference signal from the crystal clock must pass through the record/playback process of a tape recorder before being applied to one of the phase comparator inputs. The low-pass filter and voltage-controlled oscillator simulate a large flywheel that stores the average value of the playback frequency. By applying the average value to the second phase comparator input, the phase comparator output will consist of only the short-term variations from the average speed. These variations are divided into various frequency bands for further analysis. The metering circuit provides a convenient quantitative measurement of the speed variations.

Just as the sampling rate of a digital audio system determines the highest possible audio frequency that can be encoded, the frequency of the test tone determines the range of flutter components that can be measured by any frequency demodulator. The typical upper frequency is about 0.4 times the test frequency. Due to the nature of the sidebands that are required to operate the demodulator, a typical 18 kHz audio bandwidth can support a 12.5 kHz test tone and a flutter bandwidth of 5 kHz. This measurement technique, referred to as *high-band flutter measurement*, is supported by Audio Precision.

Unfortunately, most flutter meters use a low-frequency test tone of 3150 Hz and cut off all flutter components above 250 Hz, ignoring many flutter components caused by modern-day servo systems and virtually all scrape components due to the elastic vibration of the tape. To make matters worse, most flutter specifications are made through a flutter weighting filter that only measures flutter components near 4 Hz. Proper maintenance requires that a broader spectrum test be implemented to check for any possible problem.

Two methods of specifying flutter performance are commonly encountered. If a flutter-free test tape is available, the flutter reading obtained in the playback mode can be reported. Most professional recorders, however, have flutter levels that are equal to or better than any available test tapes. In this case, recording and

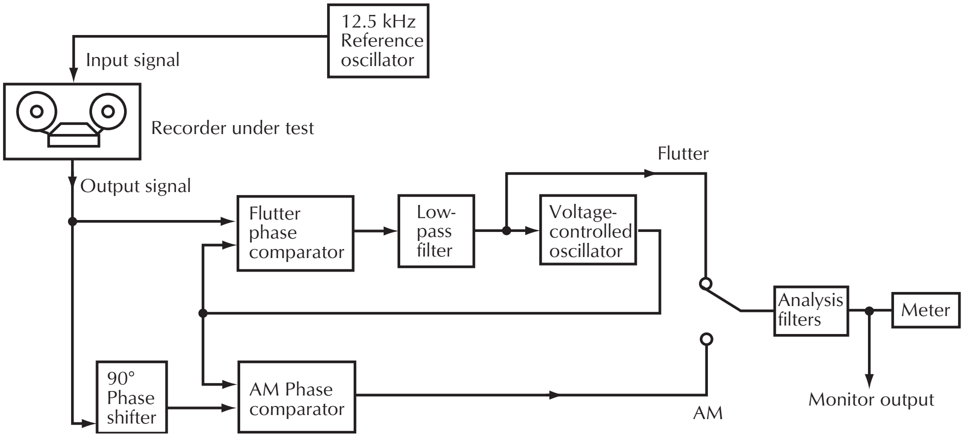


Figure 28-60. Flutter meter block diagram.

reproducing on the same machine is appropriate. The method of testing should be noted as part of the performance report.

Although test and diagnostic work is commonly conducted with simultaneous record/reproduce, the final testing should always be conducted in the reproduce-only mode. The tape should be started and stopped several times, with the various transport elements reoriented by hand between runs, to achieve a sampling of random combinations of the various record and playback flutter components. The arithmetic average of the maximum values of each sample throughout the reel, excluding any infrequent short-duration bursts, is the reported value.

If the flutter readings are excessive, the next step is to analyze the flutter waveform for information to help pinpoint which tape path component is defective. The following techniques are helpful in isolating the culprit:

1. The human ear and brain form a very versatile spectrum analyzer that frequently can immediately identify the defective component from the characteristics of the flutter signal being reproduced in a monitor loudspeaker. Take advantage of this free portable instrument that is always at your disposal by listening to the demodulated output from the flutter meter.
2. The various selectable filters of the flutter meter can be used to isolate the general portion of the flutter spectrum in which the offending component is generating flutter.
3. The expected rotational flutter rate from a rotating component can be calculated from the diameter of the component and the tape speed using the expression

   $$\text{Flutter frequency} = \frac{S_T}{\pi d} \qquad (28\text{-}13)$$

   where,
   S_T is the tape speed,
   d is the diameter of the component.

   These frequencies can range from approximately 0.5 Hz for the once-around of full reel of tape to 60 Hz for a small-diameter capstan shaft. Some manufacturers include a table of these flutter frequencies in their maintenance manuals. The small balls and retainer clips inside the ball bearings used in many rotating components generate additional not-so-obvious flutter components at frequencies higher than the once-around rate of the bearing.
4. If the flutter is very regular, the flutter pattern displayed on the oscilloscope can be utilized to calculate the frequency of the dominant flutter component. Any flutter components caused by ac motors or power supply ripple will remain stationary on the oscilloscope screen if the sweep triggering mode is set to line.
5. A common search technique is to deliberately create flutter by attaching a small piece of masking tape to the surface of a rotating component. The rate of the flutter blips created by the masking tape can then be compared with the unknown component to determine if the two rates are identical.
6. Note any change in the flutter spectrum when each of the auxiliary rotating components such as guides and flutter idlers is stalled. Stalling the defective component will cause the offending flutter component to cease. A notable exception to this case is the scrape flutter idler. Stalling a scrape flutter idler should usually double or triple the scrape flutter amplitude. If little or no increase is noted, the idler is not functioning properly. Check for dirty or damaged bearings that would keep the idler from spinning freely.

The following procedure describes a flutter test using a wide-bandwidth flutter meter, such as is shown in Fig. 28-61. The general technique also applies to other meters.

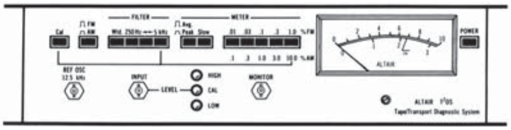


Figure 28-61. Flutter meter. Courtesy MANCO.

1. Connect the reference oscillator output (REF OSC) to the line input of the tape recorder.
2. Connect the demodulator input (INPUT) to the line output of the tape recorder.
3. Connect the demodulated output (MONITOR) to an oscilloscope and an audio monitor.
4. With the tape machine in the record mode, set the recorder's input level control to achieve a playback level of –10 VU. The green Cal light should be illuminated, indicating proper operating level.
5. With the FM/AM and Avg/Peak buttons both out and the 5 kHz and 1.0% FM buttons in, depress the Cal button. A reading of 0.68% indicates proper system operation. A 150 Hz square-wave tone will be seen on the oscilloscope and heard in the monitor.
6. To begin the actual test of the recorder, select the 250 Hz filter and choose the meter sensitivity range that yields a reading near midscale. The meter

reading is the composite value of all flutter components in the frequency band of 0.5–250 Hz, including flutter due to not only the rotating capstan, roller, and guides and their associated bearings but also any ac-power-related motor torque pulsations.

7. Select the Wtd. filter. The bandwidth is now reduced to 0.5–20 Hz, to emphasize the once-around rates due to eccentricities of the rotating components. Capstans and rollers with diameters of ½–2 in (12–50 mm) are major contributions in this band.
8. Select the 250 Hz–5 kHz bandpass filter labeled ↔. The dominant component in this range is scrape flutter, which typically peaks at 3–4 kHz for most recorders. Instabilities or oscillations of the capstan or spooling servos, which tend to occur in the 100–500 Hz range, may also be evident.
9. If the machine is equipped with a scrape flutter idler, stall the idler by pressing the point of a pencil against the top of the idler. The scrape flutter component should typically rise to two or three times the normal value. If little or no rise or even a decrease is noted, the scrape flutter idler is not functioning properly. Clean and lubricate the idler bearings according to the manufacturer's instructions. Use the flutter meter to obtain optimum positioning of the idler after cleaning.
10. Select the 5 kHz filter. This overall reading covers the entire range from 0.5 Hz–5 kHz.

28.7 Tape Testing

Contrary to popular belief, not all tape that reaches the customer's hands is fault free. Although the tape manufacturers are to be commended for the very high standards of excellence that are maintained, the customer must be prepared to deal with the bad rolls of tape that slip through the manufacturer's quality control screening. The problems that do arise can usually be traced to one of the seven steps in the manufacturing process:

1. The basic recipe of approximately a dozen major ingredients that form the oxide mixture must be correctly formulated. Each ingredient must be pure and must be measured correctly. Errors in mixing and experimental formula modifications often lead to nondurable oxides that shed debris onto the guides and heads.
2. The mixing of the ingredients must be thorough but not excessive. Inadequate mixing leads to high modulation noise and high background noise. Excessive mixing reduces noise but increases print through.
3. The coating process must apply a uniform coating across the width and length of the tape. The coating is applied to jumbo rolls that range from 18–36 inch (0.5–1 m) in width. To monitor the entire width of one of these rolls fully would require over 400 channels of conventional record/reproduce circuits!
4. The tape is baked to remove solvents by passing the coated web through a multizone oven. Poor temperature control can lead to either brittle or soft oxides.
5. The jumbo roll is run through heated rollers that make the oxide denser to increase output and high-frequency response. This calendaring step is a major factor in determining the modulation noise content of the finished tape.
6. The tape is slit to the final width by a set of rotary shears. Poor slitting can produce ruffled edges, wavy or crooked tape, and excessive oxide and backing debris on the recording surface.
7. The tape is rewound onto reels or hubs, tested, and then packaged for sale. The tape cartons usually pass through a very large degausser so that no residual signals are left on the tape.

Mistakes during the manufacturing process create four types of problems. The most common of these is signal amplitude variations, which are due to either a nonhomogeneous magnetic dispersion or erratic tape-to-head contact due to physical distortions of the tape. Other common problems include excessive noise or distortion and high print through.

A common method of testing the signal instability and dropouts is to observe the amplitude variations of a sine-wave signal on either an oscilloscope or a VU meter. While these techniques give some insight into the performance of the tape, they do not yield a quantitative value that can be used for determining acceptable limits of performance.

A more informative method is to amplitude demodulate the test signal to remove the steady tone and magnify the fluctuations. If the output of the demodulator is properly filtered and fed to a metering circuit, quantitative values for the fluctuations in various test bandwidths can be read.

Unlike other flutter test instruments, the flutter meter shown in Fig. 28-49 contains amplitude-demodulating circuitry to be used for testing tape. The AM test configuration is identical to the previous flutter setup, except that the FM/AM selector is set for AM mode testing to connect the phase-lock loop as a synchronous amplitude demodulator. The AM meter ranges, which are ten times larger than the flutter ranges, are labeled below the meter ranging pushbuttons.

The AM reading for 15 in/s (38 cm/s) operation is typically 0.5% rms for a good roll of tape on a professional recorder. The texture of the demodulation products coming from the audio monitor should be a low rumbling with only occasional moderate bursts. The high-pass filter ↔ should produce a uniform hiss.

Typical symptoms of bad rolls of tape include readings that are approximately three times higher than the normal readings or very large frequent bursts that drive the meter pointer hard against the upper stop. Routine studio tests of large quantities of tape stock over a period of two years has shown that these easily spotted characteristics are good indicators of defective tape.

Although amplitude variations are symptomatic of bad tape, the tape transport and heads are also possible sources. If the tape is not being held snugly against the faces of the heads due to inadequate tape tension, the tape may suffer irregular spacing loss. Other contributors are dirt on the heads or heads that have been worn so flat that the gap is no longer pressed firmly against the tape. Mechanical misalignments, such as a twisted head or improperly positioned guides or scrape flutter idlers, can also degrade the contact between the tape and head.

Misadjustments of the bias amplitude or even-order distortions of the bias or erase waveforms can also produce excessive AM levels. Always verify that the bias levels and tuning are correct before condemning the tape.

A simple method of avoiding embarrassment when a defective roll of tape is suspected is to recheck the machine with a reference roll of the same type of tape that is known to be good. If changing from the reference roll to the suspect roll causes a large increase in AM content, then the tape is the source of the problem.

Since none of the tape manufacturers supplies information that is useful for specifying the AM performance of a tape, the user must generate data by testing several rolls of tape on machines. Once this process is begun, subsequent additions to the database will provide even more insight into the expected range of values.

28.8 Magnetic Head Troubleshooting and Maintenance

Troubleshooting any piece of complex equipment requires a methodical search technique to isolate the source of the problem quickly. The most productive technique is to conduct a series of tests that subdivide the faulty portion of the total system into smaller and smaller parts until the fault source is finally isolated.

Applying this technique to a magnetic tape recorder would lead to partitioning questions such as:

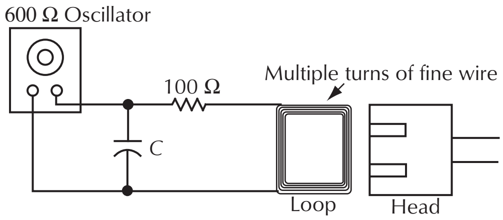
1. Is the problem associated with the tape drive, the audio circuitry, or the control logic?
2. Does the fault occur during recording, playback, and/or input monitoring?
3. Is the problem due to the recorder or the roll of tape?
4. Is the problem similar at both tape speeds?
5. Is the problem the same throughout the reel of tape?
6. Does temperature or running time have an effect?

If the problem relates to the audio signal passing through the recorder, a fundamental question that must be answered is whether the problem is wavelength-dependent or frequency-dependent. Wavelength problems immediately isolate the problem to the interface between the moving tape and the heads. Frequency problems are often related to the audio circuits.

A very useful tool for separating wavelength problems from frequency problems is a simple device known as a *flux loop* shown in Fig. 28-62. The flux loop, which consists of nothing more than a few turns of fine magnet wire driven with a constant current from an audio oscillator, creates a magnetic field that simulates a perfect lossless piece of tape. When the flux loop is attached to the gap region of the playback head, the flux from the loop excites the head much like the primary winding on a transformer excites the secondary winding. This direct excitation eliminates all the wavelength effects associated with gap length, azimuth error, and thickness loss. If the reproduce electronics perform correctly when excited by the flux loop but still fail to reproduce a known-good prerecorded test tape correctly, the problem is a wavelength-dependent error at the head-to-tape interface.

The playback response from a simple flux loop is by no means flat. Since the dominant loss due to the coating thickness is not present for flux loop excitation, the high-frequency response with a flux loop will show a pronounced rise that relates to the particular reproduce equalization standard that is being utilized. NAB low-frequency equalization will also produce a roll-off below 50 Hz.

To simplify the measurement process, the oscillator signal feeding the flux loop is usually preequalized to accommodate these effects of the equalization standard. Fig. 28-62 includes a simple circuit for correcting the high end, with capacitor values for several common equalizations. (The 600 Ω impedance of the oscillator is



EQ standard	$C = \frac{\text{EQ time constant}}{R_{source}\ R_{loop}}$
30 in/s	0.204 μF
1.5 in/s IEC	0.408 μF
7.5 & 15 m/s NAB	0.583 μF
7.5 in/s IEC	0.816 μF

Equalized flux loop

Figure 28-62. Flux loop.

part of the filter. For a 50 Ω oscillator, multiply the capacitor values by 2.57.) The resulting high-frequency playback response of an equalized flux loop will be flat except for any residual high-frequency discrepancies due to eddy current losses or self-resonance of the playback head and cabling.

The flux loop can also be used in reverse as a pickup device to probe the magnetic fields generated at the gaps of the record and erase heads. If the driving network is disconnected and the loop connected directly to the inputs of an oscilloscope and meter, the relative magnitude of the bias and audio fields can be examined. Care must be exercised to correct for the 6 dB/octave rise in flux loop output voltage due to the inductive nature of the flux loop. (A resistor in series with the input and a capacitor shunted across the input can be used to create an integrating low-pass filter that will flatten out this 6 dB/octave rise.)

Details regarding the construction and use of a flux loop, along with detailed mechanical alignment procedures for azimuth, height, and tape wrap, are available from the various tape recorder manufacturers.

28.8.1 Head Relapping

The performance characteristics gradually change as the abrasive action of the tape wears away the faces of the heads. The resulting decreases in gap depth will reduce shunting effects, leading to an increase in efficiency for both the record and playback heads. Bias and audio levels must be gradually reduced to offset the rising efficiency. A critical point is reached, however, when the useful face of the head has been completely removed and the length of the gap begins to increase quickly with wear. The top end of the playback response will drop abruptly within a matter of only a few hours of use, rendering the recorder unusable. At this point, the head must be replaced to restore normal performance.

The heads on most recorders require attention long before this point of ultimate failure is reached. On most machines, the tape wears away the rounded apex at the gap of the head, leading to a drop in contact pressure with the tape at the gap. The tape begins to lift off the head slightly, creating erratic short-wavelength performance due to the spacing loss effect.

The common solution is to recontour the face of the head to restore the contact pressure. This process, known as *head relapping*, can be utilized two or three times during the useful life of a head to restore original performance. Although the average technician can be trained in the relapping process, the high cost of a mistake with a 2 inch (50 mm) multitrack head assembly suggests that the more exotic relapping tasks should be handled by relapping specialists.

28.9 Routine Signal Alignment Procedure

A common problem arises with conventional recorders and alignment procedures—namely, that the procedures require a change in each adjustment to verify that the optimum point has been reached. This typically leads to not only the premature demise of many trimmer potentiometers (which are typically rated by the manufacturer for a life of 200 adjustment cycles) and head azimuth hardware, but also many operator errors due to the tedious nature of adjusting a multitrack machine that may have as many as 1000 adjustments.

If the operator is willing to adopt a philosophy that most of the adjustments are probably adequately close to optimum and that they need not be readjusted, then the alignment task shifts to looking for the exceptions to the norm rather than arbitrarily resetting everything. This strategy promotes better results since each iteration of the alignment procedure serves to fine-tune the results rather than to erase all past efforts and start afresh for each alignment with a high probability of error.

A few exceptions to the need for tweaking to verify proper performance are worthy of note. For example, head azimuth can be verified with a differential method that uses alternating test segments that have equal but opposite amounts of deliberate azimuth error. If the drop in level is equal for both directions of tilt, then the head must be correctly aligned to the correct vertical reference. No head adjustments are required if the test results are satisfactory.

A similar noninvasive test procedure for optimizing the bias level can be achieved if the bias system contains a master bias level trimmer that varies the level of bias for all tracks simultaneously. The bias level can be increased and decreased on all tracks with this single control to verify that the proper level of overbias is achieved without resorting to unnecessary adjustments on each track.

The following sequence of steps represents a comprehensive alignment procedure that would be appropriate whenever the proper performance of a recorder must be verified. Since the details of each step vary with machine type, the operator should consult the operator's manual published by the manufacturer of the recorder.:

1. Clean and inspect the tape transport. (Refer to Section 28.8).
2. Degauss the heads and guides. (Refer to Section 28.3.8). Before using a degausser, always verify that the tips of the unit are covered with a soft material such as plastic or tape that will not scratch the faces of the magnetic heads.
3. Calibrate the reproduce section of the recorder with a test tape of known accuracy. Several brands of standard alignment tapes are available for this purpose. Remember that the final results will be no better than the measurement standard that is being used as a reference.

   First, verify the perpendicular alignment of the reproduce head with the short-wavelength azimuth test tone on the test tape. The azimuth and/or phase alignment of the head can be measured with an oscilloscope using either a Lissajous pattern or a dual-trace display or with a phase meter that reads phase error directly. If no specialized equipment is available, invert one channel and sum the inverted output with another channel that is not inverted. Phase alignment produces a deep null in the summed output. Since phase alignment at one frequency does not eliminate the possibility of a 360° error, check the phase for several lower frequencies. The voice announcements on the alignment tapes provide a convenient multifrequency sample for this purpose.

   Next, establish a convenient reference level for making playback frequency-response measurements. Check and adjust the high-frequency reproduce equalizer at 10 kHz to match this reference level. Once the equalizer has been set, sweep through the tones on the tape, noting the maximum deviations from the reference value. Readjust the equalizer and the reference level as necessary to obtain the desired degree of flatness.

   When the results are satisfactory, write down the results for later comparison. Having a record of correct performance makes troubleshooting much easier.

   Two pitfalls exist when making the previously discussed adjustments: one affects the reference level and the other affects the frequency- response and reference level. Some recorders use different track widths for the record and playback heads. For machines that have wider playback heads, the full-track test tapes used for most of the wide-tape formats will produce an enhanced output during testing. The reference level from the tape must be set above the 0 VU reference by the amount of this extra pickup due to the wider head when using the playback head. When setting the reference level for sync/overdub playback, the track width is correct, yielding a true 0 VU level that requires no correction.

   If the record head has a wider track, then the normal playback level will be correct and the error will occur on the sync/overdub level.

   The second problem is created by the fringing effect of long wavelengths that produces a rise in playback response at low frequencies whenever additional flux is present beyond the area being scanned by the reproduce head. Such a condition exists for playback of a full-track alignment tape and for test and alignment procedures that apply the same low-frequency signal to all tracks of the recorder simultaneously.

   The fringing effect will first create a problem in establishing the correct reference level for the midband-level set tone. At 15 in/s and 30 in/s (38 cm/s and 76 cm/s) tape speeds, sufficient fringing may exist to create an error of approximately 0.5–1 dB, depending on the track format, tape speed, and geometry of the head cores and shielding. This extra fringing contribution in the reference tone also makes the high-frequency response appear to be deficient, tempting the operator to raise the equalizer adjustment. Consult the operator's manual for the correct procedure and correction factors for a given model of recorder.

   The final step in the reproduce alignment procedure is to set the level and equalization of the sync/overdub circuit. The operator may choose to defer the azimuth alignment of the record head until the following record alignment procedure if the heads have not been disturbed.
4. The record alignment begins with the verification and/or adjustment of the azimuth setting of the record head. Using the playback head as a standard, set the record head alignment while recording a

short-wavelength signal such as a 10 kHz or 15 kHz signal to give minimum azimuth or phase error using whatever method was used for the reproduce alignment procedure. This alignment should be rechecked after the bias and record equalization settings are made, since these adjustments can introduce varying amounts of phase delay.

The bias should be set by adjusting for the desired amount of overbias as recommended by the tape and machine manufacturer for the appropriate type of tape, record head gap width, and tape speed. Note that a 10 kHz signal at 30 in/s (76 cm/s) does not achieve the desired wavelength of 1.5 mils (38 µm) that is typically specified for bias adjustment. The test frequency must be changed to match the tape speed.

The bias should first be decreased to achieve deliberate underbias, and then slowly increased to the point at which a peak in the playback level is observed. Continue to increase the bias until the signal drops by the number of decibels desired. Typical overbias settings range from 2–5 dB for professional formats.

Once the bias is correctly adjusted, the input signal should be set to the frequency used as a reference during the playback alignment. The record gain control can then be set to produce the reference level when driven with the appropriate 0 VU input level.

Adjust the high-frequency record equalizer to match the record/play response as closely as possible to the alignment tape response noted previously. Smoothness in the midband frequencies is more important than trying to hold small errors at 15 kHz or 20 kHz.

Recheck the record head azimuth to verify that changes in bias and equalization have not created any phase differences. Readjust as necessary until all parameters are optimized.

Set the record gain preset and the input monitor gain calibration to achieve a 0 VU reading in all monitor modes.

5. After the record section has been aligned, a final test and alignment of the low-frequency playback equalizers can be undertaken. To eliminate all the fringing problems previously mentioned, the equalizers should be set in the record/play mode with signal being applied to every other track. Make small adjustments as required to optimize the smoothness of the response.

   If any large discrepancies are noticed, rerun the alignment tape. Any failure in the low-frequency record equalizer circuits, such as a faulty switching component, will create an error that should be obvious if a large correction is required. If any doubt still exists, record a full-frequency sweep and then flip the reels over to play the tape backward. The alignment should be similar within a few tenths of a dB to the values set in the forward direction.

6. The alignment procedure is not completed until the noise level and erasure have been checked. Record a signal at +6 VU, rewind the tape, and then erase the signal. Listen on the monitor speakers to the level of the residual signal and to the subjective nature of the tape noise. The tone should be either completely eliminated or well buried in the tape noise. The noise should be a smooth hiss without large or frequent bursts or crackling. All tracks should be similar in performance. Also, check for objectional clicks and pops when changing modes.

   Although these noise and erasure levels can be read from instruments, the operator should take the time to listen to the machine before issuing his or her stamp of approval. Many sessions have died aborning because the recorder was never given a final listening test after alignment.

The previous procedure does not include several steps that are more appropriately considered to be maintenance routines. Examples include tuning of the bias and erase sources, tuning of bias traps, checking meter calibration, and testing distortion levels. These tests are not required on a day-to-day basis.

As a final note on alignment, never gloss over large discrepancies. The corrections that should be required for this alignment procedure should be on the order of a small part of a dB, not several dB. Whenever a large change seems required, stop long enough to determine why such a large change is necessary. Look for faulty components and recheck your own procedure. Recheck the maintenance log to establish the proper level of performance that should be expected. Heeding the small symptoms may help you avoid a serious catastrophic failure.

28.10 Automated Alignment

The onslaught of digital technology has provided the tools to control the variable alignment adjustments of a tape recorder with a microprocessor. Multiple sets of calibration constants can be stored in nonvolatile memory, permitting rapid changes of operating speeds, equalization standards, reference flux levels, and tape types.

Once the provisions for automated adjustment are made available, three methods of alignment are possible. Under the simplest mode, the operator performs a manual alignment with the calibration constants being stored for later use. This method permits rapids changeovers, but does not simplify bias and equalization adjustments to optimize a specific roll of tape.

If the microprocessor can be provided with input information from the metering devices on the individual tracks, then calibration programs can be automatically executed without operator intervention. The program contains the "strategy" for alignment, including desired amounts of overbias, equalization adjustment frequencies, and operating levels. Beware that such systems use an inferred adjustment technique which does not actually test many of the critical parameters. For example, the recorder will set the bias level for minimum distortion based on an overbias criterion at a specified frequency. In reality, the machine doesn't have the ability to measure distortion. The strategy only infers that overbiasing by the desired amount corresponds to minimum distortion. Unfortunately, if a malfunction exists that causes abnormal operation, the adjustment routine may not detect the symptoms.

Nearly automatic calibration can be implemented by connecting external automated test equipment such as an Audio Precision System One test set to the machine through an external intelligent controller such as an IBM-compatible computer. A remotely controlled input/output switching matrix will also be necessary for multitrack machines. An operator is still required to adjust nonautomated devices such as head azimuth and to change tape reels for calibration tapes and sample stock. The calibration program of the intelligent controller sequences through a comprehensive set of tests which rigorously exercise the machine. Parameters such as harmonic and intermodulation distortions, crosstalk, erasure, flutter, speed, noise, and phase can be tested against absolute standards of acceptance.

Hopefully, the advent of inexpensive DSP (digital signal processor) chips will allow manufacturers to include the diagnostic equipment as a part of the built-in calibration hardware.

A final word of caution is appropriate at this point. Many operators and test technicians ignore symptoms that indicate problems are developing in a tape recorder. A good example is the frequent need to boost the high-frequency equalization adjustments of a recorder. A properly operating machine should not show such trends, but a gradually deteriorating head would create just such a problem. Simply readjusting without determining the cause of the change wastes an opportunity to fix a problem at an early stage before it grows to catastrophic consequences. Try to avoid problems by fixing things before they break completely.

Chapter 29

MIDI

by David Miles Huber

29.1 Introduction to MIDI 1101
  29.1.1 What MIDI Isn’t 1101
29.2 The MIDI Message 1101
  29.2.1 MIDI Channels 1102
  29.2.2 MIDI Modes 1103
    47.2.2.1 Poly/Mono 1104
    47.2.2.2 Omni On/Off 1104
  29.2.3 Channel Messages 1104
  29.2.4 System Messages. 1108
29.3 Hardware Systems within MIDI Production 1110
  29.3.1 The MIDI Interface 1113
  29.3.2 Hardware and Software Electronic Instruments 1114
  29.3.3 Performance and Parameter Controllers 1117
29.4 Sequencers 1118
  29.4.1 A Basic Introduction to Sequencers 1119
  29.4.2 Other Software Sequencing Applications 1121
29.5 Multimedia and the Web 1125
  29.5.1 Standard MIDI Files 1125
  29.5.2 General MIDI 1126
29.6 MIDI-Based Synchronization 1126
  29.6.1 MIDI Real-Time Messages 1127
  29.6.2 Song Position Pointer 1128
  29.6.3 MIDI Time Code 1128

This page intentionally left blank

29.1 Introduction to MIDI

Simply stated, Musical Instrument Digital Interface (MIDI) is a digital communications language and compatible specification that allows multiple hardware and software electronic instruments, performance controllers, computers, and other related devices to communicate with each other over a connected network. MIDI is used to translate performance- or control-related events (such as playing a keyboard, selecting a patch number, varying a modulation wheel, triggering a staged visual effect, etc.) into equivalent digital messages and then transmit these messages to other MIDI devices where they can be used to control sound generators and other performance parameters. The beauty of MIDI is that its data can be easily recorded into a hardware device or software program (known as a sequencer), where it can be edited and transmitted to electronic instruments or other devices to create music or control any number of parameters.

In artistic terms, this digital language is an important medium that lets artists express themselves with a degree of flexibility and control that wasn't possible at an individual level beforehand. Through the use of this performance language, an electronic musician can create and develop a song or composition in a practical, flexible, affordable, and fun production environment.

The word *interface* refers to the actual data communications link and software/hardware systems in a connected MIDI network. Through MIDI, it's possible for all of the electronic instruments and devices within a network to communicate real-time performance and control-related MIDI data messages throughout a system to multiple instruments and devices via MIDI, USB, or FireWire networked data lines. Given that MIDI data can simultaneously transmit performance and control messages over multiple channels (usually in groupings of 16 channels per port), an electronic musician can record, overdub, mix, and play back their performances in a building-block fashion that resembles the multitrack recording process. In fact, the true power of MIDI lies in its ability to edit, control, alter and automate parts of a composition after the original performance has been recorded, allowing performance parameters to be easily altered in ways that are unique to the medium.

29.1.1 What MIDI Isn't

For starters, let's dispel one of MIDI's greatest myths: MIDI doesn't communicate audio it cannot create sounds! It is a digital language protocol that can only be used to trigger and/or control a device (which, in turn generates, reproduces, or controls the sound). Thus, the MIDI data and the audio routing paths are kept entirely separate from each another, Fig. 29-1. Even if they digitally share the same transmission cable (such as through USB or FireWire), the actual data paths and formats are distinct.

In short, MIDI's control-related language can be thought of as the dots on a player-piano roll—when we put the paper roll up to our ears, we hear nothing. However, when the cutout dots pass over the sensors on a player piano, the instrument itself begins to make beautiful music. The analogy is pretty much the same with MIDI. A MIDI file or data stream is simply a set of instructions that pass through wires in a serial fashion, but when an electronic instrument interprets the data, we then hear sound.

As a performance-based control language, MIDI complements modern music production, by allowing a performance track to be edited, layered, altered, spindled, mutilated, and improved with relative ease under completely automated computer control and after the fact, during post-production. If you played a bad note, fix it. If you want to change the key or tempo of a piece, change it. If you want to change the expressive volume of a phrase in a song, just do it! Even its sonic character (timbre) can be changed! These capabilities merely hint at the power of this medium that widely affects the project studio, professional studio, audio or visual and film, live performance, multimedia, and even your cell phone!

29.2 The MIDI Message

From its inception in the early 80s, the MIDI 1.0 spec (which is still the adopted version to this day) must be strictly adhered to by those who design and manufacture MIDI-equipped instruments and devices. As such, users needn't worry about whether the MIDI Out of one device will be understood by the MIDI In of a device that's made by another manufacturer (at least the basic performance level). We need only consider the day-to-day dealings that go hand-in-hand with using electronic instruments, without having to be concerned with data compatibility between devices.

MIDI messages are communicated through a standard MIDI line in a serial fashion at a speed of 31,250 bits/s. These messages are made up of groups of 8-bit words (known as bytes), which are used to convey instructions to one or all MIDI devices within a system. Only two types of bytes are defined by the MIDI specification: the status byte and the data byte.

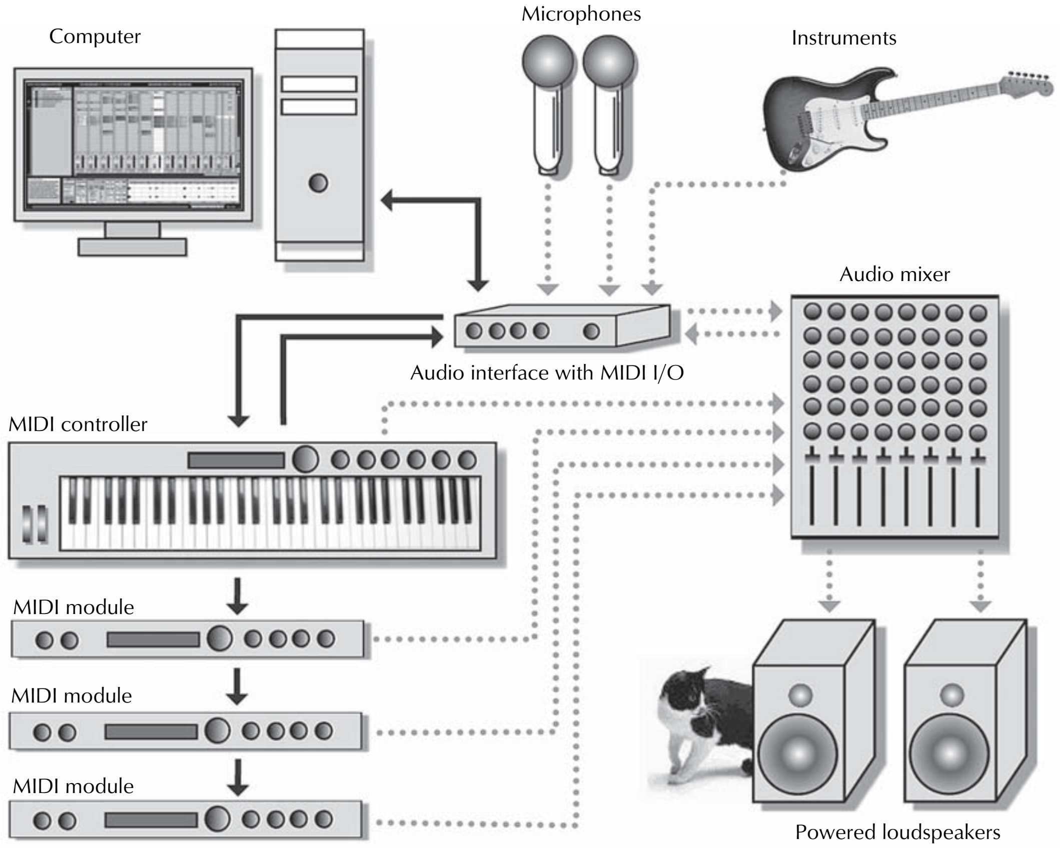


Figure 29-1. Example of a typical MIDI system with the MIDI network connections being shown in solid lines and audio connections shown using dotted lines.

A status byte is used to identify what type of MIDI function is to be performed by a device or program. It's also used to encode channel data (allowing the instruction to be received by a device that's set to respond to a specific channel). A Data byte is used to associate a value to the event that's given by the accompanying status byte.

The most significant bit (MSB), the leftmost binary bit within a digital word within a MIDI byte, is used solely to identify the data's particular function. The MSB of a status byte is always 1, while the MSB of a data byte is always 0. For example, a 3 byte MIDI note-on message (which is used to signal the beginning of a MIDI note) in binary form might read as shown in Table 29-1. Thus, a 3 byte note-on message of (10010100) (01000000) (01011001) will transmit instructions that would be read as "Transmitting a note-on message over MIDI channel #5, using keynote #64, with an attack velocity (volume level of a note) of 89."

Table 29-1. Status and Data Byte Interpretation

	Status Byte	**Data Byte 1**	**Data Byte 2**
Description	Status/Channel #	Note #	Attack Velocity
Binary Data	(1001.0100)	(0100.0000)	(0101.1001)
Numeric Value	(Note On/Ch #5)	(64)	(89)

29.2.1 MIDI Channels

Just as a public speaker might single out and communicate a message to one individual in a crowd, MIDI messages can be directed to communicate information to a specific device or series of devices within a MIDI system. This is done by imbedding a channel-related nibble (4 bits) within the status byte, allowing data to be conveyed to any of 16 channels over a single MIDI data cable line, Fig. 29-2. This makes it possible for performance or control information to be communicated to a

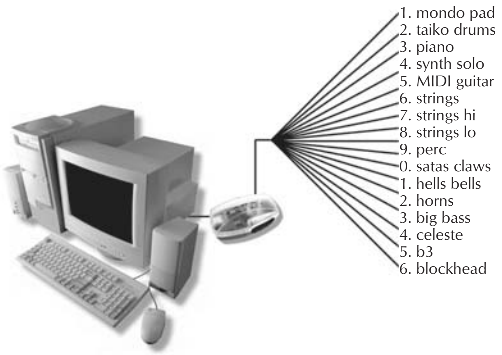


Figure 29-2. Up to 16 channels can be transmitted through a single MIDI cable.

0000 = CH#1	0100 = CH#5	1000 = CH#9	1100 = CH#13
0001 = CH#2	0101 = CH#6	1001 = CH#10	1101 = CH#14
0010 = CH#3	0110 = CH#7	1010 = CH#11	1110 = CH#15
0011 = CH#4	0111 = CH#8	1011 = CH#12	1111 = CH#16

specific device or a sound generator within a device that's assigned to a particular channel.

Whenever a MIDI device, sound generator, or program function is instructed to respond to a specific channel number, it will only respond to messages that are transmitted on that channel (i.e., it ignores channel messages that are transmitted on any other channel). For example, let's assume that we're going to create a short song using a synthesizer that has a built-in sequencer (a device or program that's capable of recording, editing, and playing back MIDI data) and two other synths, Fig. 29-3.

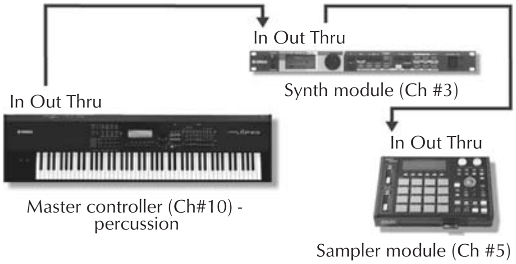


Figure 29-3. MIDI setup showing a set of MIDI channel assignments.

1. We could start off by recording a drum track into the master synth using channel 10 (many synths are pre-assigned to output drum/percussion sounds on this channel).
2. Once recorded, the sequence will then transmit the notes and data over channel 10, allowing the synth's percussion section to be heard.
3. Next, we could set a synth module to channel 3, and instruct the master synth to transmit on the same channel (since the synth module is set to respond to data on channel 3, its generators will sound whenever the master keyboard is played). We can now begin recording a melody line into the sequencer's next track.
4. Playing back the sequence will then transmit data to both the master synth (percussion section) and the module (melody line) over their respective channels. At this point, our song is beginning to take shape.
5. Now, we can set a sampler (or other instrument type) to respond to channel 5, and instruct the master synth to transmit on the same channel, allowing us to further embellish the song.
6. Now that the song's complete, the sequencer can then play the musical parts to the synths on their respective MIDI channels, all in an environment that allows us to have complete control of volume, edit, and a wide range of functions over each instrument. In short, we've created a true multichannel working environment.

It goes without saying that the above example is just but one of the infinite setup and channel possibilities that can be encountered in a production environment. It's often true, however, that even the most complex MIDI and production rooms will have a system, a basic channel and overall layout that makes the day-to-day operation of making music easier. This layout and the basic decisions in your own room are, of course, up to you. Streamlining a system to work both efficiently and easily will come over time with experience and practice.

29.2.2 MIDI Modes

Electronic instruments often vary in the number of sounds and/or notes that can be simultaneously produced by their internal sound-generating circuitry. For example, certain instruments can only produce one note at a single time (known as a *monophonic* instrument), while others can generate 16, 32, and even 64 notes at once (these are known as *polyphonic* instruments). The latter type is easily capable of playing chords and/or more than one musical line on a single instrument.

In addition, some instruments are only capable of producing a single generated sound patch (often referred to a voice) at any one time. Its generating circuitry could be polyphonic, allowing the player to lay down chords and bass/melody lines), but it can only produce these notes using a single, characteristic sound at any one time (e.g., an electric piano, or a synth bass, or a string patch). However, the vast majority of newer synths differs from this in that they're multi-timbral in nature, meaning that they can generate numerous sound patches at any one time (e.g., an electric piano, and a synth bass, and a string patch). That is it's common to run across electronic instruments that can simultaneously generate a number of voices, each offering its own control over parameters (such as volume, panning, modulation, etc.) and—best of all—it's also common for different sounds to be assigned to their own MIDI channels, allowing multiple patches to be internally mixed within the device (often top a stereo output bus), or to independent outputs.

As a result of these differences between instruments and devices, a defined set of guidelines (known as *MIDI reception modes*) has been specified that allows a MIDI instrument to transmit or respond to MIDI channel messages in several ways. For example, one instrument might be programmed to respond to all 16 MIDI channels at one time, while another might be polyphonic in nature, with each voice being programmed to respond to only a single MIDI channel.

47.2.2.1 Poly/Mono

An instrument or device can be set to respond to MIDI data in either the poly mode or the mono mode. Stated simply, an instrument that's set to respond to MIDI data polyphonically will be able to play more than one note at a time. Conversely, an instrument that's set to respond to MIDI data monophonically will only be able to play a single note at any one time.

47.2.2.2 Omni On/Off

Omni on/off refers to how a MIDI instrument will respond to MIDI messages at its input. When Omni is turned on, the MIDI device will respond to all channel messages that are being received regardless of its MIDI channel assignment. When Omni is turned off, the device will only respond to a single MIDI channel or set of assigned channels (in the case of a multitimbral instrument).

The following list and figures explain the four modes that are supported by the MIDI spec in more detail.

- **Mode 1—Omni On/Poly:** In this mode, an instrument will respond to data that's being received on any MIDI channel, and then redirect this data to the instrument's base channel, Fig. 29-2A. In essence, the device will play back everything that's presented at its input in a polyphonic fashion… regardless of the incoming channel designations. As you might guess, this mode is rarely used.
- **Mode 2—Omni On/Mono:** As in Mode 1, an instrument will respond to all data that's being received at its input, without regard to channel designations. However, this device will only be able to play one note at a time, Fig. 29-2B. Mode 2 is used even more rarely than Mode 1, as the device can't discriminate channel designations and can only play one note at a time.
- **Mode 3—Omni Off/Poly:** In this mode, an instrument will only respond to data that matches its assigned base channel in a polyphonic fashion, Fig. 29-2C). Data that is assigned to any other channel will be ignored. This mode is by far the most commonly used because as it allows the voices within a multi-timbral instrument to be individually controlled by messages that are being received on different MIDI channels. For example, each of the 16 channels in a MIDI line could be used to independently play each of the parts in a 16-voice, multitimbral synth.
- **Mode 4—Omni Off/Mono:** As with Mode 3, an instrument will be able to respond to performance data that's transmitted over a single, dedicated channel; however, each voice will only be able to generate one MIDI note at a time, Fig. 29-2D. A practical example of this mode is often used in MIDI guitar systems, where MIDI data is monophonically transmitted over six consecutive channels (one channel/voice per string).

29.2.3 Channel Messages

Channel-voice messages are used to transmit real-time performance data throughout a connected MIDI system. They're generated whenever a MIDI instrument's controller is played, selected, or varied by the performer. Examples of such control changes could be the playing of a keyboard, pressing of program selection buttons, or movement of modulation or pitch wheels. Each channel-voice message contains a MIDI channel number within its status byte, meaning that only devices that are assigned to the same channel number will respond to these commands. There are seven channel-voice message types: note-on, note-off, polyphonic-key pressure, channel pressure, program change, pitch-bend change and control change.

Note-On Messages. A note-on message is used to indicate the beginning of a MIDI note. It is generated each time a note is triggered on a keyboard, controller, or other MIDI instrument (i.e., by pressing a key, hitting a drum pad, or by playing a sequence).

A Note-On message consists of 3 bytes of information, Fig. 29-4.

Note-on status/MIDI channel number, MIDI pitch number and Attack velocity value.

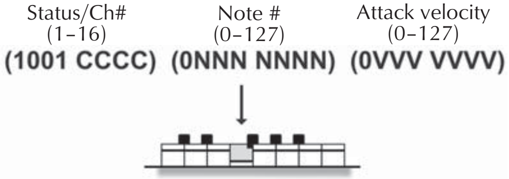


Figure 29-4. Byte structure of a MIDI note-on message.

The first byte in the message specifies a note-on event and a MIDI channel (1–16). The second byte is used to specify which of the possible 128 notes (numbered 0–127) will be sounded by an instrument. In general, MIDI note number 60 is assigned to the middle C key of an equally tempered keyboard, while notes 21 to 108 correspond to the 88 keys of an extended keyboard controller. The final byte is used to indicate the velocity or speed at which the key was pressed (over a value range that varies from 0 to 127). Velocity is used to denote the loudness of a sounding note, which increases in volume with higher velocity values (although velocity can also be programmed to work in conjunction with other parameters such as expression, control over timbre, sample voice assignments, etc).

Note-Off Messages. A note-off message is used as a command to stop playing a specific MIDI note. Each note-on message will continue to play until a corresponding note-off message for that note has been received. In this way, the bare basics of a musical composition can be encoded as a series of MIDI note-on and note-off events. It should also be pointed out that a note-off message wouldn't cut off a sound; it'll merely stop playing it. If the patch being played has a release (or final decay) slope, it will begin this stage upon receiving the message.

A note-off message consists of three bytes of information, Fig. 29-5. Note-off status/MIDI channel number, MIDI pitch number and Attack velocity value.

In contrast to the dynamics of attack velocity, the release velocity value (0–127) indicates the velocity or speed at which the key was released. A low value indicates that the key was released very slowly, whereas a high value shows that the key was released quickly. Although not all instruments generate or respond to MIDI's release velocity feature, instruments that are capable of responding to these values can be programmed to vary a note's speed of decay, often reducing the signal's decay time as the release velocity value is increased.

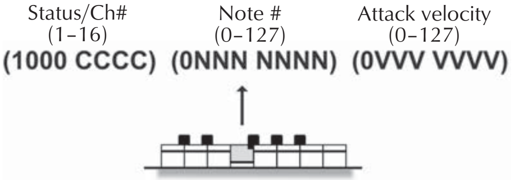


Figure 29-5. Byte structure of a MIDI note-off message.

A note-on message that contains an attack velocity of 0 (zero) is generally equivalent to the transmission of a note-off message. This common implementation tells the device to silence a currently sounding note by playing it with a velocity (volume) level of 0.

All Notes Off. On the odd occasion (often when you least expect it), a MIDI note can get stuck! This can happen when data drops out or a cable gets disconnected, creating a situation where a note receives a note-on message, but not a note-off message, resulting in a note that continues to plaaaaaaaaaaayyyyyyyyyy! Since you're often too annoyed or under pressure to take the time to track down which note is the offending sucka… it's generally far easier to transmit an all notes off message that silences everything on all channels and ports. If it exists, this can easily be done by pressing a Panic Button that's built into the sequencer or hardware MIDI interface.

Pressure (Aftertouch) Messages. Pressure-related messages (often referred to as *aftertouch*) occur after you've pressed a key and then decide to press down harder to gain a particular effect. For devices that can respond to (and therefore generally transmit) these messages, aftertouch can often be assigned to such parameters as vibrato, loudness, filter cutoff, and pitch. Two types of pressure messages are defined by the MIDI spec:

- Channel-pressure.
- Polyphonic-pey pressure.

Channel-pressure messages are commonly transmitted by instruments that only respond to a single overall pressure, regardless of the number of keys that are being played at any one time, Fig. 29- 6. For

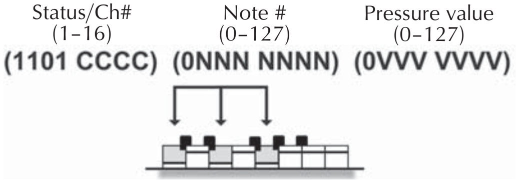


Figure 29-6. Byte structure of a MIDI channel-pressure message.

example, if six notes are played on a keyboard and additional aftertouch pressure is applied to just one key, the assigned parameter would be applied to all six notes.

A channel-pressure message consists of 3 bytes of information, Fig. 29-6: Channel-pressure status/MIDI channel number, MIDI note number, and pressure value.

Polyphonic-key pressure messages respond to pressure changes that are applied to the individual keys of a keyboard. That's to say that a suitably equipped instrument can transmit or respond to individual pressure messages for each key that's depressed.

How a device responds to these messages will often vary from manufacturer to manufacturer (or can be assigned by the user). However, pressure values are commonly assigned to such performance parameters as vibrato, loudness, timbre, and pitch. Although controllers that are capable of producing polyphonic pressure are generally more expensive, it's not uncommon for an instrument to respond to these messages.

A polyphonic-key pressure message consists of 3 bytes of information, Fig. 29-7. Polyphonic-key pressure status/MIDI channel number, MIDI note number, and pressure value.

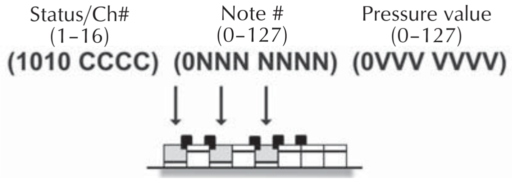


Figure 29-7. Byte structure of a MIDI polyphonic-key pressure message.

Program-Change Messages. Program-change messages are used to change a MIDI instrument or device's active program or preset number. A preset is a user- or factory-defined number that actively selects a specific sound patch or system setup. Using this extremely handy message, up to 128 presets can be remotely selected from another device or controller. For example:

- A program-change message can be transmitted from a remote keyboard or controller to an instrument, allowing sound patches to be remotely switched, Figs. 29-8 and 29-9.
- Program-change messages could be programmed at the beginning of a sequence, so as to instruct the various instruments or voice generators to set to the correct sound patch before playing.
- It could be used to alter patches on an effects device, either in the studio or on stage. The list goes on.

A program-change message, Fig. 29-8, consists of 2 bytes of information: program-change status/ MIDI channel number and program ID number.

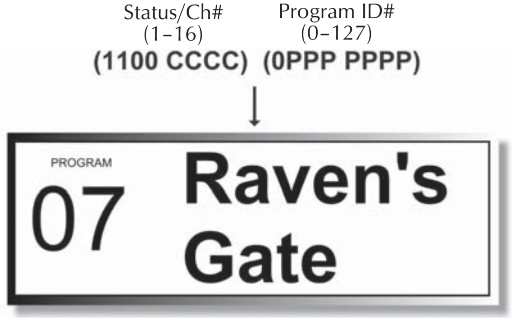


Figure 29-8. Byte structure of a MIDI program-change message.

Pitch-bend Messages. Pitch-bend sensitivity refers to the response sensitivity (in semitones) of a pitch-bend wheel or other pitch-bend controlle, which, as you'd expect, is used to bend the pitch of a note upward or downward. Since the ear can be extremely sensitive to changes in pitch, this control parameter is encoded using 2 data bytes, yielding a total of 16,384 steps. Since this parameter is most commonly affected by varying a pitch wheel, Fig. 29-9, the control values range from –8,192 to +8,191, with 0 being the instrument's or part's unaltered pitch.

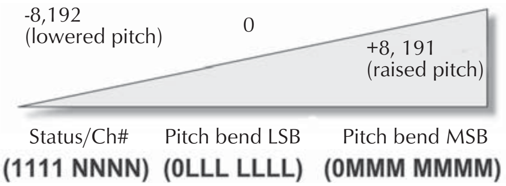


Figure 29-9. Byte structure of a pitch-bend message.

Control-Change Messages. Control-change messages are used to transmit information to a device (either internally or through a MIDI line/network) that relates to real-time control over its performance parameters.

Three types of control-change messages can be transmitted via MIDI:

1. Continuous controllers: Controllers that relay a full range of variable control settings (often ranging in value between 0–127 although, in certain cases, two controller messages can be combined in tandem to achieve a greater resolution).
2. Switch controllers: Controllers that have either an off or an on state with no intermediate settings.
3. Channel-mode message controllers: The final set of control change messages ranges between controller numbers 120 through 127, and are used to set the note sounding status, instrument reset, local control on/off, all notes off, and MIDI mode status of a device or instrument.

A single control-change message or a stream of such messages is transmitted whenever controllers (such as foot switches, foot pedals, pitch-bend wheels, modulation wheels, breath controllers, etc.) are varied in real time. Newer controllers and software editors often offer up a wide range of switched and variable controllers, allowing for extensive, user-programmable control over any number of device, voice, and mixing parameters in real-time, Fig. 29-10.

A control-change message, Fig. 29-11, consists of 3 bytes of information: control-change status/MIDI channel number, controller ID number, and corresponding controller value.

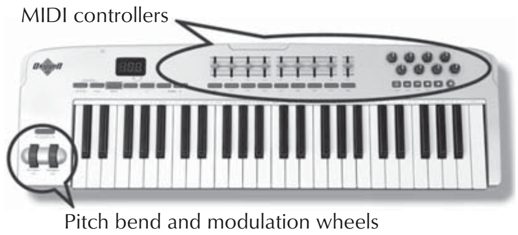


Figure 47-10. M-audio controller. Courtesy of M-Audio, a division of Avid Technology, Inc., www.m-audio.com.

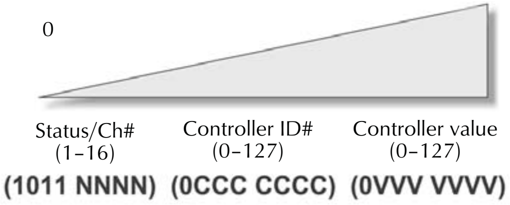


Figure 29-11. Byte structure of a control-change message.

As you can see, the second byte of the control-change message is used to denote the controller ID number. This all-important value is used to specify which of the device's program or performance parameters are to be addressed.

Table 29-2 details the general categories and conventions for assigning controller numbers to an associated parameter (as specified by the 1995 update of the MMA (MIDI Manufacturers Association, www.midi.org). This is definitely an important section to earmark, as these numbers will be an important guide towards knowing and/or finding the right ID number that can help you on your path towards finding that perfect variable for making it sound right.

Table 29-2. Listing of Controller ID Numbers, Outlining Both the Defined Format and Convention and Controller Assignments

Control Number	Parameter
14 Bit Controllers Coarse/MSB (most significant bit)	
0	Bank Select 0–127 MSB
1	Modulation Wheel or Lever 0–127 MSB
2	Breath Controller 0–127 MSB
3	Undefined 0–127 MSB
4	Foot Controller 0–127 MSB
5	Portamento Time 0–127 MSB
6	Data Entry MSB 0–127 MSB
7	Channel Volume (formerly Main Volume) 0–127 MSB
8	Balance 0–127 MSB
9	Undefined 0–127 MSB
10	Pan 0–127 MSB
11	Expression Controller 0–127 MSB
12	Effect Control 1 0–127 MSB
13	Effect Control 2 0–127 MSB
14	Undefined 0–127 MSB
15	Undefined 0–127 MSB
16–19	General Purpose Controllers 1–4 0–127 MSB
20–31	Undefined 0–127 MSB
14-bit Controllers Fine/LSB (least significant bit)	
32	LSB for Control 0 (Bank Select) 0–127 LSB
33	LSB for Control 1 (Modulation Wheel or Lever) 0–127 LSB
34	LSB for Control 2 (Breath Controller) 0–127 LSB
35	LSB for Control 3 (Undefined) 0–127 LSB
36	LSB for Control 4 (Foot Controller) 0–127 LSB
37	LSB for Control 5 (Portamento Time) 0–127 LSB

Table 29-2. Listing of Controller ID Numbers, Outlining Both the Defined Format and Convention and Controller Assignments (Continued)

Control Number	Parameter
38	LSB for Control 6 (Data Entry) 0–127 LSB
39	LSB for Control 7 (Channel Volume, formerly Main Volume) 0–127 LSB
40	LSB for Control 8 (Balance) 0–127 LSB
41	LSB for Control 9 (Undefined) 0–127 LSB
42	LSB for Control 10 (Pan) 0–127 LSB
43	LSB for Control 11 (Expression Controller) 0–127 LSB
44	LSB for Control 12 (Effect control 1) 0–127 LSB
45	LSB for Control 13 (Effect control 2) 0–127 LSB
46–47	LSB for Control 14–15 (Undefined) 0–127 LSB
48–51	LSB for Control 16–19 (General Purpose Controllers 1–4) 0–127 LSB
52–63	LSB for Control 20–31 (Undefined) 0–127 LSB
7-bit Controllers	
64	Damper Pedal On/Off (Sustain) <63 off, >64 on
65	Portamento On/Off <63 off, >64 on
66	Sustenuto On/Off <63 off, >64 on
67	Soft Pedal On/Off <63 off, >64 on
68	Legato Footswitch <63 Normal, >64 Legato
69	Hold 2 <63 off, >64 on
70	Sound Controller 1 (Default: Sound Variation) 0–127 LSB
71	Sound Controller 2 (Default: Timbre/Harmonic Intens.) 0–127 LSB
72	Sound Controller 3 (Default: Release Time) 0–127 LSB
73	Sound Controller 4 (Default: Attack Time) 0–127 LSB
74	Sound Controller 5 (Default: Brightness) 0–127 LSB
75	Sound Controller 6 (Default: Decay Time—see MMA RP-021) 0–127 LSB
76	Sound Controller 7 (Default: Vibrato Rate—see MMA RP-021) 0–127 LSB
77	Sound Controller 8 (Default: Vibrato Depth—see MMA RP–021) 0–127 LSB
78	Sound Controller 9 (Default: Vibrato Delay—see MMA RP-021) 0–127 LSB
79	Sound Controller 10 (Default undefined—see MMA RP-021) 0–127 LSB
80–83	General Purpose Controller 5–8 0–127 LSB
84	Portamento Control 0–127 LSB
85–90	Undefined
91	Effects 1 Depth (Default: Reverb Send Level) 0–127 LSB
92	Effects 2 Depth (Default: tremolo Level) 0–127 LSB
93	Effects 3 Depth (Default: Chorus Send Level) 0–127 LSB
94	Effects 4 Depth (Default: Celeste [Detune] Depth) 0–127 LSB
95	Effects 5 Depth (Default: Phaser Depth) 0–127 LSB
Parameter Value Controllers	
96	Data Increment (Data Entry +1)
97	Data Decrement (Data Entry −1)
98	Non-Registered Parameter Number (NRPN)—LSB 0–127 LSB
99	Non-Registered Parameter Number (NRPN)—MSB 0–127 MSB
100	Registered Parameter Number (RPN)–LSB* 0–127 LSB
101	Registered Parameter Number (RPN)–MSB* 0–127 MSB
102–119	Undefined
Reserved for Channel Mode Messages	
120	All Sound Off 0
121	Reset All Controllers
122	Local Control On/Off 0 off, 127 on
123	All Notes Off
124	Omni Mode Off (+ all notes off)
125	Omni Mode On (+ all notes off)
126	Poly Mode On/Off (+ all notes off)
127	Poly Mode On (+ mono off +all notes off)

The third byte of the control-change message is used to denote the controller's actual data value. This value is used to specify the position, depth, or level of a parameter. Here are a few examples as to how these values can be implemented to vary control and mix parameters.

In certain cases, greater resolutions than can be given by a single 7-bit course message (128 steps) might be available to increase a controller's resolution. This is simply accomplished by adding an additional fine controller value message to the data stream, resulting in an overall resolution that yields an overall total of 16,384 discrete steps!

29.2.4 System Messages

System Messages. As the name implies, system messages are globally transmitted to every MIDI device in the MIDI chain. This is accomplished because MIDI channel numbers aren't addressed within the byte structure of a system message. Thus, any device will respond to these messages, regardless of its MIDI channel assignment. The three system message types are system-common messages, system real-time messages, and system-exclusive messages.

System-Common Messages. System-common messages are used to transmit MIDI time code, song position pointer, song select, tune request, and end-of-exclusive data messages throughout the MIDI system or 16 channels of a specified MIDI port.

MTC Quarter-Frame Messages. MIDI time code (MTC) provides a cost effective and easily implemented way to translate SMPTE (a standardized synchronization time code) into an equivalent code that conforms to the MIDI 1.0 spec. It allows time-based codes and commands to be distributed throughout the MIDI chain in a cheap, stable, and easy-to-implement way. MTC quarter-frame messages are transmitted and recognized by MIDI devices that can understand and execute MTC commands.

A grouping of eight quarter frames is used to denote a complete time code address (in hours, minutes, seconds, and frames), allowing the SMPTE address to be updated every two frames. Each quarter-frame message contains 2 bytes. The first is a quarter-frame common header, while the second byte contains a 4-bit nibble that represents the message number (0–7). A final nibble is used to encode the time field (in hours, minutes, seconds, or frames).

Song Position Pointer Messages. As with MIDI time code, song position pointer (SPP) lets you synchronize a sequencer, tape recorder, or drum machine to an external source from any measure position within a song. The SPP message is used to reference a location point in a MIDI sequence (in measures) to a matching location within an external device. This message provides a timing reference that increments once for every six MIDI clock messages (with respect to the beginning of a composition).

Unlike MTC (which provides the system with a universal address location point), SPP's timing reference can change with tempo variations, often requiring that a special tempo map be calculated in order to maintain synchronization. Because of this fact, SPP is used far less often than MIDI time code.

Song Select Messages. Song select messages are used to request a specific song from a drum machine or sequencer (as identified by its song ID number). Once selected, the song will thereafter respond to MIDI start, stop, and continue messages.

Tune Request Messages. The tune request message is used to request that a MIDI instrument initiate its internal tuning routine (if so equipped).

End-of-Exclusive Messages. The transmission of an end-of-exclusive (EOX) message is used to indicate the end of a system-exclusive message. In-depth coverage of system-exclusive messages will be discussed later in this chapter.

System Real-Time Messages. Single-byte system real-time messages provide the all-important timing element required to synchronize all of the MIDI devices in a connected system. To avoid timing delays, the MIDI specification allows system real-time messages to be inserted at any point in the data stream, even between other MIDI messages.

Timing-Clock Messages. The MIDI timing-clock message is transmitted within the MIDI data stream at various resolution rates. It is used to synchronize the internal timing clocks of each MIDI device within the system and is transmitted in both the start and stop modes at the currently defined tempo rate.

In the early days of MIDI, these rates (which are measured in pulses per quarter note, ppq) ranged from 24 to 128 ppq. However, continued advances in technology have brought these rates up to 240, 480, or even 960 ppq.

Start Messages. Upon receipt of a timing-clock message, the MIDI start command instructs all connected MIDI devices to begin playing from their internal sequences initial start point. Should a program be in midsequence, the start command will reposition the sequence back to its beginning, at which point it will begin to play.

Stop Messages. Upon receipt of a MIDI stop command, all devices within the system will stop playing at their current position point.

Continue Messages. After receiving a MIDI stop command, a MIDI continue message will instruct all connected devices to resume playing their internal sequences from the precise point at which it was stopped.

Active-Sensing Messages. When in the stop mode, an optional active-sensing message can be transmitted throughout the MIDI data stream every 300 milliseconds. This instructs devices that can recognize this message that they're still connected to an active MIDI data stream.

System-Reset Messages. A system-reset message is manually transmitted in order to reset a MIDI device or instrument back to its initial power-up default settings (commonly mode 1, local control on, and all notes off).

System-Exclusive Messages. The system-exclusive (SysEx) message allows MIDI manufacturers, programmers and designers to communicate customized MIDI messages between MIDI devices. It's the purpose of these messages to give manufacturers, programmers, and designers the freedom to communicate any device-specific data of an unrestricted length, as they see fit. In practice, SysEx data is commonly used to communicate real-time controller information (i.e., a remote controller surface will commonly use SysEx to communicate data to/from a MIDI-capable hard- or software device. SysEx can also be used transmit and receive device-specific program, patch parameter and sample data from one instrument or device to another. For example, SysEx can be used to transmit patch and overall setup data between identical make and (most-often) model of synthesizer. Let's say that you have a Brand X Model Z synthesizer and it turns out that you have a buddy across town who also has a Brand X Model Z. That's cool, except your buddy's synth has a completely different set of sound patches that was loaded into her instrument and you want them! SysEx to the rescue! All you need to do is go over and transfer your buddy's patch data into your synth, or into a MIDI sequencer as a SysEx data dump. In order to make life easier, make sure you take your instruction manual along, (just in case you run into a snag), and follow these simple guidelines. I'll caution you that you're taking on these tasks at your own risk. Take your time; be patient and be careful during these procedures:

1. Back up your present patch data! This can be done by transmitting a SysEx dump of your synthesizer's entire patch and setup data to your sequencer's SysEx dump utility, or SysEx track on your sequencer (of course, you should get out both the device's manual and your sequencer's manual and follow their SysEx dump instructions very carefully during the process). This is so important that I'll say it again: Back up your present patch data before attempting a SysEx dump! If you forget and download a new SysEx dump, your previous settings could easily be lost.
2. Save the data, according to your sequencer's manual.
3. Check that the dump was successful by reloading it back into the device in question. Did it reload properly? If so, your current patch data is now saved.
4. Next, connect your buddy's device to your sequencer. Dump this data to your sequencer. Save the new patch data (using a new and easily identifiable file-name), according to your sequencer's manual and then safely back this data up.
5. Reconnect the sequencer to your synth and load the new data dump into it. Does your synth have a bunch of new sounds? Now reload your original SysEx dump back into your device. Are the original sounds restored?

The transmission format of a SysEx message, Fig. 29-12, as defined by the MIDI standard includes a SysEx status header, manufacturer's ID number, any number of SysEx data bytes, and an EOX byte. On receiving a SysEx message, the identification number is read by a MIDI device to determine whether or not the following messages are relevant. This is easily accomplished, because a unique 1- or 3-byte ID number is assigned to each registered MIDI manufacturer. If this number doesn't match the receiving MIDI device, the ensuing data bytes will be ignored. Once a valid stream of SysEx data is transmitted, a final EOX message is sent, after which the device will again begin responding to incoming MIDI performance messages. A detailed practical explanation of the many uses (and wonders) of SysEx can be found in the synthesizer section of Chapter 4, as well as in the patch editor section of Chapter 6. I definitely recommend that you check these out, because SysEx is one of the most cost-effective and powerful tools that an electronic musician can have. It's definitely well worth the reading!

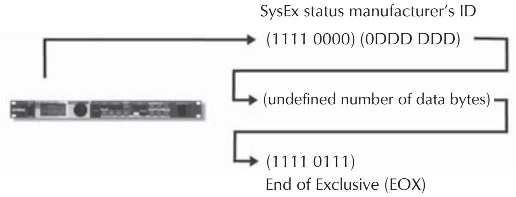
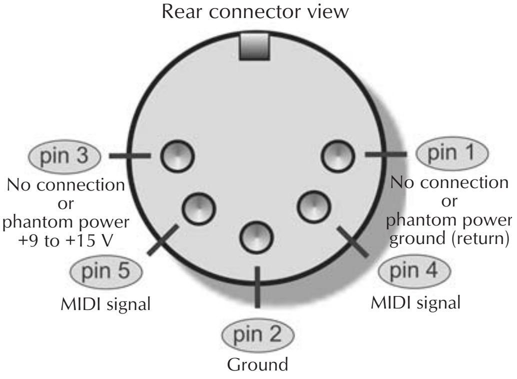

SysEx status manufacturer's ID
(1111 0000) (0DDD DDD)
(undefined number of data bytes)
(1111 0111)
End of Exclusive (EOX)

Figure 29-12. System-exclusive data (one ID byte format).

29.3 Hardware Systems within MIDI Production

As a data transmission medium, MIDI is relatively unique in the world of sound production in that it's able to pack 16 discrete channels of performance, controller, and timing information and transmit it in one direction, using data densities that are economically small and easy to manage. In this way, it's possible for MIDI messages to be communicated from a specific source (such as a keyboard or MIDI sequencer) to any number of devices within a connected network over a single MIDI data chain. In addition, MIDI is flexible enough that multiple MIDI data lines can be used to interconnect devices in a wide range of possible system configurations (for example, multiple MIDI lines can be used to transmit data to instruments and devices over 32, 48, 128, or more discrete MIDI channels!

The MIDI Cable. A MIDI cable, Fig. 29-13, consists of a shielded, twisted pair of conductor wires that has a male 5-pin DIN plug located at each of its ends. The MIDI specification currently uses only 3 of the 5 pins, with pins 4 and 5 being used as conductors for MIDI data, while pin 2 is used to connect the cable's shield to equipment ground. Pins 1 and 3 are currently not in use, although the next section describes an ingenious system for power devices through these pins, using a system that's known as MIDI phantom power. The cables themselves use twisted cable and metal shield groundings to reduce outside interference, such as radio-frequency interference (RFI) or electrostatic interference, both of which can serve to distort or disrupt the transmission of MIDI messages.


Rear connector view
pin 3
No connection or phantom power +9 to +15 V
pin 1
No connection or phantom power ground (return)
pin 5
MIDI signal
pin 4
MIDI signal
pin 2
Ground

A. Connector wiring diagram

B. Standard length MIDI cable

Figure 29-13. The MIDI cable.

MIDI cables come prefabricated in lengths of 2, 6, 10, 20, and 50 feet, and can commonly be obtained from music stores that specialize in MIDI equipment. To reduce signal degradations and external interference that tends to occur over extended cable runs, 50 feet is the maximum length specified by the MIDI specification. (As an insider tip, I found that Radio Shack is also a great source for picking up 3 and 6 feet MIDI cables at a fraction of what you'd sometimes spend at a music store).

MIDI Phantom Power. In December 1989, Craig Anderton wrote an article in *Electronic Musician* about a proposed idea for allowing a source to provide a standardized 12 Vdc power supply to instruments and MIDI devices directly through pins 1 and 3 of a basic MIDI cable. Although pins 1 and 3 are technically reserved for possible changes in future MIDI applications, over the years several forward-thinking manufacturers (and project enthusiasts) have begun to implement MIDI phantom power directly into their studio and on-stage systems.

Wireless MIDI. In recent times, a number of companies have begun to manufacturer wireless MIDI transmitters that can allow a battery-operated MIDI guitar, wind controller, etc. to be footloose and fancy free on-stage and in the studio. Working at distances of up to 500 feet, these battery-powered transmitter/receiver systems introduce very low delay latencies and can be switched over a number of radio channel frequencies.

MIDI Jacks. MIDI is distributed from device to device using three types of MIDI jacks: MIDI In, MIDI Out, and MIDI Thru, Fig. 29-14. These three connectors use 5-pin DIN jacks as a way to connect MIDI Instruments, devices, and computers into a music and/or production network system. As a side note, it's nice to know that these ports (as strictly defined by MIDI 1.0 Spec.) are optically isolated to eliminate possible ground loops that might occur when connecting numerous devices together.

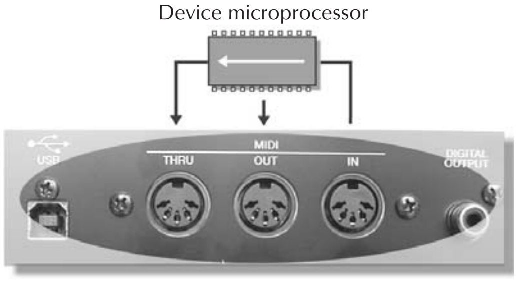


Figure 29-14. MIDI in, out, and thru ports, showing the device's signal path routing.

- **MIDI In**—The MIDI In jack receives messages from an external source and communicates this performance, control, and/or timing data to the device's internal microprocessor, allowing an instrument to be played and/or a device to be controlled. More than one MIDI In jack can be designed into a system to provide for MIDI merging functions or for devices that can support more than 16 channels (such as a MIDI Interface). Other devices (such as a controller) might not have a MIDI In jack at all.
- **MIDI Out**—The MIDI Out jack is used to transmit MIDI performance, control messages or SysEx from one device to another MIDI instrument or device. More than one MIDI Out jack can be designed into a system, giving it the advantage of controlling and distributing data over multiple MIDI paths using more than just 16 channels (i.e., 16 channels × *N* MIDI port paths).
- **MIDI Thru**—The MIDI Thru jack retransmits an exact copy of the data that's being received at the MIDI In jack. This process is important, because it allows data to pass directly through an instrument or device to the next device in the MIDI chain. Keep in mind that this jack is used to relay an exact copy of the MIDI In data stream and isn't merged with data being transmitted from the MIDI Out jack.
- **MIDI Echo**—Certain MIDI devices may not include a MIDI Thru jack, at all. Certain of these devices, however, may give the option of switching the MIDI Out between being an actual MIDI Out jack and a MIDI Echo jack, Fig. 29-15. As with the MIDI Thru jack, a MIDI echo option can be used to retransmit an exact copy of any information that's received at the MIDI In port and route this data to the MIDI Out/Echo jack. Unlike a dedicated MIDI Out jack, the MIDI Echo function can often be selected to merge incoming data with performance data that's being generated by the device itself. In this way, more than one controller can be placed in a MIDI system at one time. It should be noted that although performance and timing data can be echoed to a MIDI Out/Echo jack, not all devices can echo SysEx data.

Typical Configurations. Although electronic studio production equipment and setups are rarely alike (or even similar), there are a number of general rules that make it easy for MIDI devices to be connected into a functional network. These common configurations allow MIDI data to be distributed in the most efficient and understandable manner possible.

As a primary rule, there are only two valid ways to connect one MIDI device to another within a MIDI chain, Fig. 29-16:

1. Connecting the MIDI Out jack of a source device (controller or sequencer/computer) to the MIDI In of a second device in the chain.
2. Connecting the MIDI Thru jack of the second device to the MIDI In jack of the third device in the chain and following this same Thru-to-In convention until the end of the chain is reached.

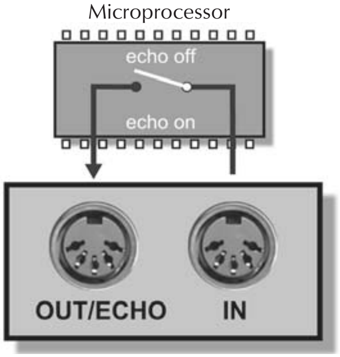


Figure 29-15. MIDI echo configuration.

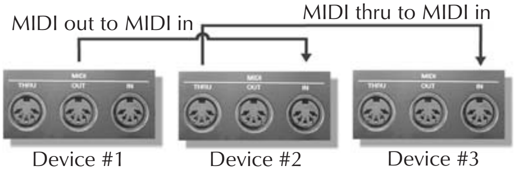
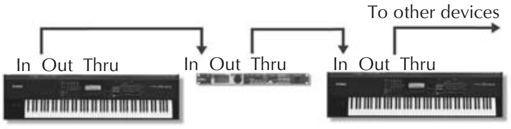


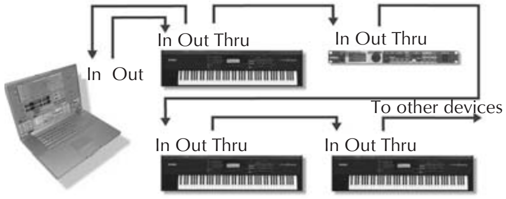
Figure 29-16. The two valid means of connecting one MIDI device to another.

The Daisy Chain. One of the simplest and most common ways to distribute data throughout a MIDI system is the daisy chain. This method relays MIDI data from a source device (controller or sequencer/computer) to the MIDI In jack of the next device in the chain (which receives and acts upon this data). In turn, this device relays an exact copy of this incoming data out to its MIDI Thru jack, which is then relayed to the next device in the chain. This device can then relay an exact copy of this incoming data out to its MIDI Thru jack, which is then relayed to the next device in the chain… etc. In this way, up to 16 channels of MIDI data can be chained from one device to the next within a connected data network—and it's precisely this concept of transmitting multiple channels through a single MIDI line that makes this concept work! Let's try to understand this concept better by looking at a few examples.

Fig. 29-17A shows a simple (and common) example of a MIDI daisy chain, whereby data flows from a controller (MIDI Out jack of the source device) to a synth module (MIDI In jack of the second device in the chain), where an exact copy of this data is relayed from its MIDI Thru jack to another synth (MIDI In jack of the third device in the chain). From the section on MIDI channels in Chapter 2, it shouldn't be hard to understand that if our controller is transmitting on MIDI channel 2, the second synth in the chain (which is set to channel 2) will ignore the messages and not play while the 3rd synth (which is set to channel 3) will be playing its heart out. The moral of this story is that although there's only one connected data line, a wide range of instruments and channel voices can be played in a surprisingly large number of combinations, all by using individual channel assignments along a daisy chain.



A. Typical daisy chain hookup.



B. Example of how a computer can be connected into a daisy chain.

Figure 29-17. Example of a connected MIDI system using a daisy chain.

Another example, Fig. 29-17B, shows how a computer can easily be designated as the master source within a daisy chain, so that a sequencing program could be used to control the entire playback and channel routing functions of a daisy-chained system. In this situation, the MIDI data flows from a master controller/synth to the MIDI In jack of a computer's MIDI interface—where the data can be played into, processed, and rechannelized through a MIDI sequencer. The MIDI Out of the interface is then routed back to the MIDI In jack of the master controller/synth (which receives and acts on this data). In turn, the controller relays an exact copy of this incoming data out to its MIDI Thru jack, which is then relayed to the next device in the chain. This device can then relay an exact copy of this incoming data out to its MIDI Thru jack, which is then relayed to the next device in the chain etc. When we stop to think about this second example, the controller is used to perform into the MIDI sequencer, which then is used to communicate this edited and processed performance data out to the various instruments throughout the connected MIDI chain.

The Multiport Network. Another common approach to routing MIDI throughout a production system involves distributing MIDI data through the multiple 2, 4 and 8 In/Out ports that are available on a newer multiport MIDI interfaces or through the use of multiple MIDI interfaces (typically these are USB devices).

In larger, more complex MIDI systems, a multiport MIDI network, Fig. 29- 17, offers several advantages over a single daisy chain path. One of the most important is its ability to address devices within a complex setup that requires more than 16 MIDI channels. For example, a 2 × 2 MIDI interface that offers up two-independent In/Out paths is capable of addressing up to 32 channels simultaneously (i.e., port A 1–16 and port B 1–16), whereas an 8 × 8 port interface is capable of addressing up to 128 individual MIDI channels.

29.3.1 The MIDI Interface

Although computers and electronic instruments both communicate using the digital language of 1s and 0s, computers simply can't understand the language of MIDI without the use of a device that translates the serial messages into a data structure that computers can comprehend. Such a device is known as the *MIDI interface*.

A wide range of MIDI interfaces currently exist that can be used with most computer systems and OS platforms. For the casual and professional musician, interfacing MIDI into a production system can be done in a number of ways. Probably the most common way to access MIDI In, Out, and Thru jacks is on a modern-day USB or FireWire audio interface or instrument/DAW controller surface. It's become a common matter for portable devices to offer 16 channels of I/O (on one port), while multi-channel interfaces often include multiple MIDI I/O ports that can give you access to 32 or more channels.

Another additional option is to choose a USB MIDI interface that can range from devices that include a single I/O port (16 channels) to a multiport system that can easily handle up to 128 channels over eight I/O ports. The multiport MIDI interface, Fig. 29-18, is often the device of choice for most professional electronic musicians who require added routing and synchronization capabilities. These rack-mountable USB devices can be used to provide eight independent MIDI Ins and Outs to easily distribute MIDI and time code data through separate lines over a connected network.

Figure 29-18. M-Audio MIDISPORT 4×4 MIDI interface. Courtesy of M-Audio, a division of Avid Technology, Inc., www.m-audio.com.

29.3.2 Hardware and Software Electronic Instruments

Since its inception in the early 80s, MIDI-based electronic musical instruments have helped to shape the face and sounds of our modern music culture. These devices (along with digital audio and advances in recording equipment technology) have altered music production, through the creation of one of the most cost-effective and powerful tools in the development of music history —the personal project studio.

The following is a sample listing of the many hardware MIDI instrument types that are currently available on the market.

The Synth. A synthesizer, Fig. 29-19, is an electronic instrument that uses multiple sound generators to create complex waveforms that can be combined (using various waveform synthesis techniques) into countless sonic variations. These synthesized sounds have become a basic staple of modern music and vary from sounding cheesy, to those that closely mimic traditional instruments all the way to those that generate rich, otherworldly sounds that literally defy classification.

Synthesizers (also known as synths) generate sounds and percussion sets using a number of different technologies or program algorithms. The earliest synths were analog in nature and generated sounds using additive or subtractive FM (frequency modulation) synthesis. This process generally involves the use of at least two signal generators (commonly referred to as *operators*) to create and modify a voice. Often, this is done through the analog or digital generation of a signal that modulates or changes the tonal and amplitude characteristics of a base carrier signal. More sophisticated FM synths can use up to 4 or 6 operators per voice and also often use filters and variable amplifier types to alter the signal's characteristics into a sonic voice that either roughly imitates acoustic instruments or creates sounds that are totally unique.

Figure 29-19. Bass Station analogue bass synth. Courtesy of Novation Digital Music Systems, Ltd.; www.novation-music.com.

Another technique that's used to create sounds is wavetable synthesis. This technique works by storing small segments of digitally sampled sound into a read-only memory chip. Various sample-based synthesis techniques use sample looping, mathematical interpolation, pitch shifting, and digital filtering to create extended and richly textured sounds that use a very small amount of sample memory.

Synthesizers are also commonly designed into rack- or half-rack-mountable modules, Fig. 29-20, that contain all of the features of a standard synthesizer, except that they don't incorporate a keyboard controller. This space-saving feature means that more synths can be placed into your system and can be controlled from a master keyboard controller or sequencer, without cluttering up the studio with redundant keyboards.

Figure 29-20. Yamaha MOTIF-RACK ES synth. Courtesy of Yamaha Corporation of America, www.yamaha.com.

Software Synthesis and Sample Resynthesis. Since wavetable synthesizers derive their sounds from prerecorded samples that are stored in a digital memory medium, it logically follows that these sounds can also be stored on hard disk (or any other medium) and loaded into the RAM memory of a personal computer. This process of downloading wavetable samples into a computer and then manipulating these samples is used to create what is known as a *virtual* or *software synthesizer*, Fig. 29-21.

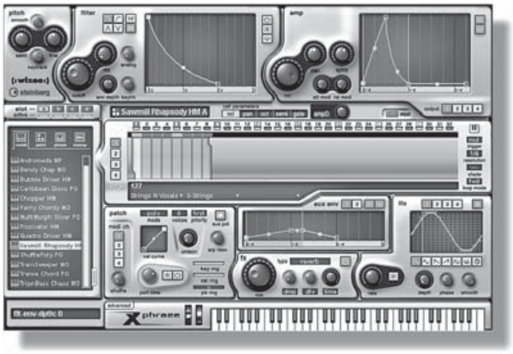


Figure 29-21. Steinberg xphrase VSTi software synth. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

In recent years, software synths have grown from being novel and obscure programs that were primarily used by the academic community to their present state of being widely accepted in the production community as a cost-effective musical instrument. These software modules can be used in conjunction with a digital audio workstation to offer up a wide range of complex sounds that can mimic traditional instruments, as well as create sonic textures that are both new and interesting.

Sample resynthesis software systems are able to take software synthesis to a new level, by allowing the user to build, save, and recall sonic patches that can be built from traditional synthesis building blocks (such as oscillators, voltage-controlled amplifiers, voltage-controlled filters, and mixers). In addition to sound generation, digital audio samples can be imported and re-synthesized in a way that can create sounds of almost any texture or type that you can possibly imagine. All of these software blocks can be combined in a graphic environment that allows these instruments, textures, and soundscapes to be easily saved to disk for later recall.

Using various internal software data communications protocols, it's possible to communicate MIDI, audio, timing sync and control data between an instrument (or effect plug-in) and a host DAW program/CPU processor. These plug-in protocols make it possible for much or all of the audio and timing data to be routed through the host audio application, allowing the instrument or application to either integrate into the DAW or application or to work in tandem so as to route the audio and performance/control data through the host application with relative ease. A few of these protocols include:

- Steinberg's VST (Virtual Studio Technology).
- MOTU's MAS (MOTU Audio System).
- Propellerheads ReWire.

Samplers. A sampler, Fig. 29-22, is a device that can convert audio into a digital form and/or manipulate prerecorded sampled data, using the system's own random access memory (RAM). Once loaded into RAM, the sampled audio can be edited, transposed, processed, and played in a polyphonic musical fashion.

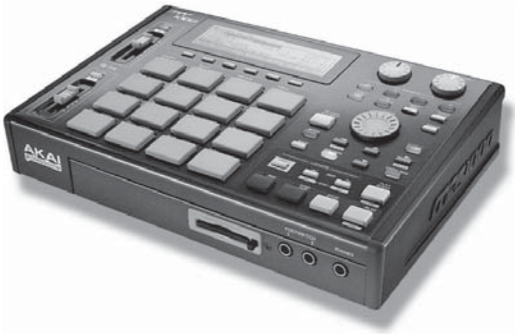


Figure 29-22. Akai MPC-1000 Music Production Center. Courtesy of Akai Professional, www.akaipro.com.

Basically, a sampler can be thought of as a wavetable synth that lets you record, load, and edit samples into RAM memory. Once loaded, these sounds (whose length and complexity are often limited only by memory size and your imagination) can be looped, modulated, filtered, and amplified (according to user or factory setup parameters), in a way that allows the waveshapes and envelopes to be modified. Signal processing capabilities, such as basic editing, looping, gain changing, reverse, sample-rate conversion, pitch change, and digital mixing capabilities can also be altered and/or varied.

A hardware sampler's design will often include a keyboard or set of trigger pads that let you polyphonically play samples as musical chords, sustain pads, triggered percussion sounds, or sound effect events. These samples can be played according to the standard Western musical scale (or any other scale, for that matter) by altering the playback sample rate over the controller's note range. For example, pressing a low-pitched key on the keyboard will cause the sample to be played back at a lower sample rate, while pressing a high-pitched one will cause the sample to be played back at rates that would put Mickey Mouse to shame. By choosing the proper sample rate ratios, sounds can be polyphonically played (whereby multiple notes are sounded at once) at pitches that correspond to standard musical chords and intervals.

A sampler (or synth) with a specific number of voices (i.e., 64 voices) simply means that up to 64 notes can be simultaneously played on a keyboard at any one time. Each sample in a multiple-voice system can be assigned across a performance keyboard, using a process known as *splitting* or *mapping*. In this way, a sound can be assigned to play across the performance surface of a controller over a range of notes, known as a *zone*, Fig. 29-23. In addition to grouping samples into various zones, velocity can enter into the equation by allowing multiple samples to be layered across the same keys of a controller, according to how soft or hard they are played. For example, a single key might be layered so that pressing the key lightly would reproduce a softly recorded sample, while pressing it harder would produce a louder sample with a sharp, percussive attack. In this way, mapping can be used to create a more realistic instrument or wild set of soundscapes that change not only with the played keys, but with velocity ranges as well.

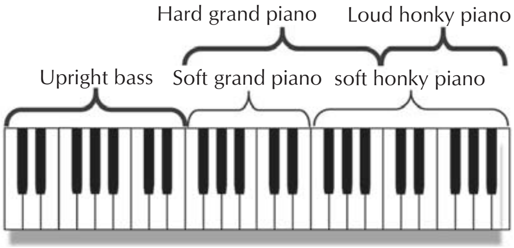


Figure 29-23. Samples can be mapped to various zones on a keyboard.

In addition to hardware sampling systems, a growing number of virtual or software samplers exist that use a computer's existing memory, processing, and signal routing capabilities in order to polyphonically reproduce samples in real time.

Offering much of the same functionality as their hardware counterparts, these software-based systems, Fig. 29-24, are capable of editing, mapping, and splitting sounds across a MIDI keyboard, using on-screen graphic controls and DAW integration that has improved to the point of equaling or surpassing their hardware counterparts in cost-effectiveness, power, and ease of use.

As with a software synth, software samplers derive their sounds from recorded and/or imported audio data that is stored as digital audio data within a personal computer. Using the DSP capabilities of today's computers (as well as the recording, sequencing, processing, mixing, and signal routing capabilities of most digital audio workstations), most software samplers are able to store and access samples within the internal memory of a laptop or desktop computer. Using a graphic interface, these sampling systems often allow the user to:

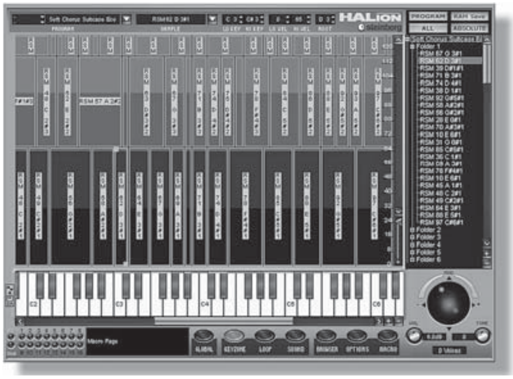


Figure 29-24. Steinberg's HALion VST software sampler. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

- Import previously recorded soundfiles (often in WAV, AIF, and other common formats)
- Edit and loop sounds into a usable form
- Vary envelope parameters (i.e., dynamics over time)
- Vary processing parameters
- Save the edited sample performance setup as a file for later recall

Software sampler systems are also often able to communicate MIDI, audio, timing sync and control data between a hard- or software instrument and a host DAW program/CPU processor, allowing for a wide range of control and setup recall.

The Drum Machine. The drum machine is most commonly a sample-based digital audio device that can't record audio into its internal memory (although this has changed in recent years, allowing it to import, record, and manipulate sampled audio much like a sampler). Traditionally, these hardware or software systems use ROM-based, prerecorded samples to reproduce high-quality drum sounds. These factory-loaded sounds often include a wide assortment of drum sets, percussion sets, rare and wacky percussion hits, and effected drum sounds (i.e., reverberated, gated, etc.). Who knows, you might even encounter "Hit me!" screams from the venerable King of Soul—James Brown.

Most hardware drum machines allow prerecorded samples to be assigned to a series of playable keypads that are often located on the machine's top face. This provides a straightforward controller surface that usually includes velocity and aftertouch dynamics. Drum voices can be assigned to each pad and edited using such control parameters as tuning, level, output assignment, and panning position. Multiple outputs are often provided, enabling individual or groups of voices to be routed to a specific output on a mixer or console.

Although a number of hardware drum machine designs include a built-in sequencer, it's more likely that these workhorses will be triggered from a MIDI sequencer. This lets us take full advantage of the real-time performance and editing capabilities that a sequencer has to offer. For example, sequenced patterns can easily be created in step time (where notes are entered and assembled into a rhythmic pattern one note at a time) and can then link together into a song that's composed of several rhythmic patterns. Alternately, performing into a sequencer on-the-fly can help create a live feel or you can combine step- and real-time tracks to create a human-sounding composite rhythm track. In the final analysis, the style and approach to composition is entirely up to you.

In addition to their hardware counterparts, an increasing number of software drum and groove instrument plug-ins have come onto the market that allow for drum patterns to be added to a production in a wide range of pattern and playing styles, Fig. 29-25.

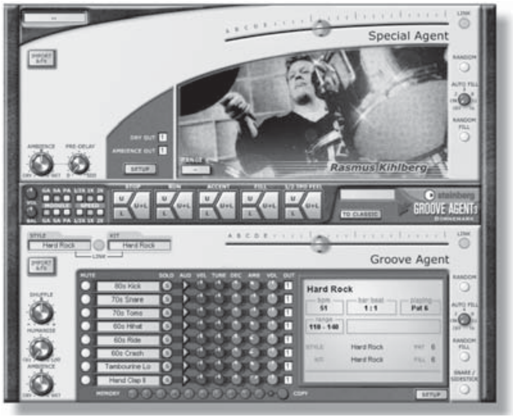


Figure 29-25. Groove Agent 3 VST Virtual Drummer. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

29.3.3 Performance and Parameter Controllers

MIDI performance controllers are used to translate the voicings and expressiveness of a musical performance into MIDI data, while a parameter controller surface is used to alter the control variables of a workstation, device or instrument.

It should be noted that a MIDI controller is expressly designed to control other devices (be they for sound, light or mechanical control) within a connected system. It contains no internal tone generators or sound-producing elements. Instead, it offers a wide range of controls for handling control, trigger and device switching events. In short, controllers have become an integral part of music production, and are available in many incarnations to control and emulate many types of musical instruments.

Keyboard Controller. The MIDI keyboard controller, Fig. 29-26, is a keyboard device that's expressly designed to control hard/software synths, samplers, modules and other devices within a connected MIDI system. It contains no internal tone generators or sound-producing elements. Instead, its design includes a performance keyboard and controls for handling MIDI performance, control, and device switching events.

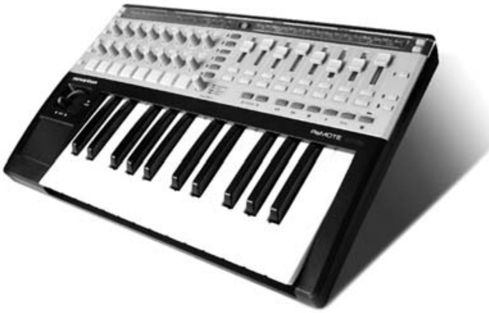


Figure 29-26. Novation ReMOTE 25SL MIDI Controller/Keyboard. Courtesy of Novation Digital Music Systems, Ltd, www.novationmusic.com.

Percussion Controllers. MIDI percussion controllers are used to translate the voicings and expressiveness of a percussion performance into MIDI data. These devices are great for capturing the feel of a live performance, while giving you the flexibility of recording and automating a performance within a DAW/sequencer environment. These controllers vary over a wide range from being a simple and cost-effective setup (i.e., using the pads on a drum machine, keys on a keyboard surface, or pads on an intro-level drum controller) to a full-blown drum kit that mimics its acoustic cousin, Figs. 29-27 and 29-28.

Wind Controllers. MIDI wind controllers are expressly designed to bring the breath and key articulation of a woodwind or brass instrument to a MIDI performance. These controller types are used because many of the dynamic- and pitch-related expressions (such as breath and controlled pitch glide) simply can't be communi-

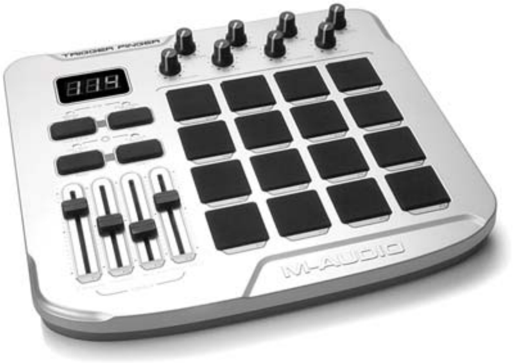


Figure 29-27. Trigger Finger 16-Pad MIDI Drum Control Surface. Courtesy of M-Audio, a division of Avid Technology, Inc., www.m-audio.com.

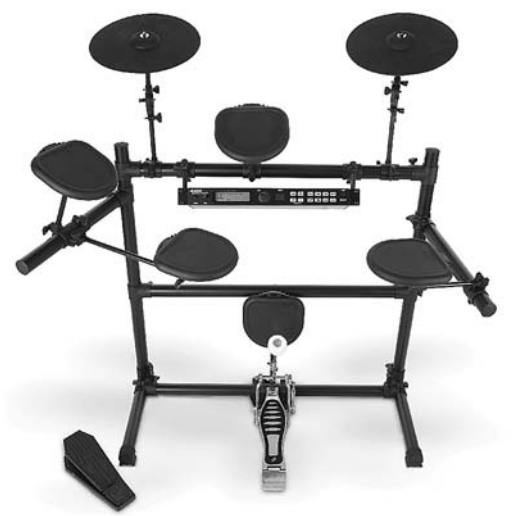


Figure 29-28. DM5 Electronic Drum Kit. Courtesy of Alesis, www.alesis.com.

cated from a standard music keyboard. In these situations, wind controllers can often help create a dynamic feel that's more in keeping with their acoustic counterparts by using an interface that provides special touch-sensitive keys, glide- and pitch-slider controls, and real-time breath sensors for controlling dynamics.

MIDI Guitars. Guitar players often work at stretching the vocabulary of their instruments beyond the traditional norm. They love doing nontraditional gymnastics using such tools of the trade as distortion, phasing, echo, feedback, etc. Due to advances in guitar pickup and microprocessor technology, it's also possible for the notes and minute inflections of guitar strings to be accurately translated into MIDI data. With this innovation, many of the capabilities that MIDI has to offer are available to the electric (and electronic) guitarist. For example, a guitar's natural sound can be layered with a synth pad that's been transposed down, giving it a rich, thick sound that just might shake your boots. Alternately, recording a sequenced guitar track into a session would give a producer the option of changing and shaping the sound later in mixdown! On-stage program changes are also a big plus for the MIDI guitar, allowing the player to radically switch between guitar voices from the guitar or sequencer or by stomping on a MIDI foot controller.

29.4 Sequencers

Apart from electronic musical instruments, one of the most important tools that can be found in the modern-day project studio is the MIDI sequencer. Basically, a sequencer is a digital device that's used to record, edit, reproduce, and distribute MIDI messages in a sequential fashion. Most sequencers function using a traditional track-based interface, separating different instruments, voices, beats, etc. in a way that makes it easier for us humans to view MIDI data as though they were linear tracks on a DAW or tape machine.

These virtual tracks contain MIDI-related performance and control events that are made up of such channel and system messages as note on/off, velocity, modulation, aftertouch, and program/continuous-controller messages. Once a performance has been recorded into a sequencer's memory, these events can be graphically (or audibly) edited into a musical performance, played back and saved to a digital storage media for recall at any time.

Integrated Sequencers. Some of the newer and more expensive keyboard synth and sampler designs include a built-in sequencer. These portable keyboard workstations have the advantage of letting you take both the instrument and sequencer on the road without having to drag a computer along.

Integrated sequencers are designed into an instrument for the sole purpose of sequencing MIDI data, and include integrated controls for performing sequence-specific functions. Ease of use and portability are often the advantages of a hardware sequencer, most of which are designed to emulate the basic functions of a tape transport (record, play, start/stop, fast forward, and rewind).

These devices generally offer a moderate amount of editing features, including note editing, velocity and other controller messages, program change, cut and paste and track merging capabilities, tempo changes, etc. Programming, track, and edit information is commonly viewed on a liquid crystal display (LCD) that's often limited in size and resolution and generally limits information to a single parameter or track at a time.

These sequencers often don't offer a wide range of editing tools beyond standard transport functions; punch-in/out commands and other basic edit tools. However, they're often more than adequate for capturing and reproducing a performance and can be integrated with other instruments that are connected in a MIDI chain.

Software Sequencers. By far, the most common sequencer type is the software MIDI sequencer. These programs or integrated components of a digital audio workstation take advantage of the versatility that a computer can offer in the way of speed, flexibility, digital signal processing, memory management, and signal routing.

Computer-based sequencers offer numerous functional advantages over their hardware counterparts. Among these are increased graphic capabilities (which often offers extensive control over track- and transport-related functions), standard computer cut and paste techniques, an on-screen graphic environment (allowing easy manipulation of program and edit-related data), routing of MIDI to multiple ports in a connected system, and the graphic assignment of instrument voices via program change messages (not to mention the ability to save and recall files using standard computer memory media). Now, let's take a look into how these devices function.

29.4.1 A Basic Introduction to Sequencers

When dealing with any type of sequencer, one of the most important concepts to grasp is the fact that these devices don't store sound directly—instead, they encode MIDI messages that instruct an instrument to play a particular note, over a certain channel, at a specific velocity and with any optional controller values. In other words, a sequencer stores music-related data commands that follow in a sequential order, which then tells instruments and/or devices how their voices are to be played and/or controlled. This simple (but important) fact means that the amount of encoded data is far less memory intensive than its hard disk audio or video recording counterparts and that the data overhead that's required by MIDI is very small. In short, a computer-based sequencer can simultaneously operate in a digital audio, digital video, processing environment without placing an additional significant load on a computer's CPU.

As you might expect, many sequencer types are currently on the market, with each offering its own set of advantages and disadvantages. It's also true that each sequencer has its own basic operating feel, and thus, choosing the best tool and toy for the job or studio is totally up to you.

Recording. From a functional standpoint, a sequencer is used as a digital workspace for creating personal compositions in environments that range from the bedroom to more elaborate project studios. Whether they're hardware or software-based, most sequencers use a working interface that's designed to emulate the traditional multitrack recording environment. A tapelike transport lets you move from one location to the next using standard Play, Stop, FF, REW and Rec command buttons. Beyond using traditional record-enable button(s) to arm selected recording track(s), all you need to do is select the MIDI input (source) and outputs (destination) ports, instrument/voice MIDI channel, instrument patch and other setup information, press the record button, and start playing.

Once you've finished laying down a track, you can jump back to any point in the sequence and listen to your original track while continuing to lay down additional MIDI tracks until the song begins to form.

Almost all sequencers are capable of punching in and out of record while playing a sequence. This common function lets you drop in and out of record on a track (or tracks) in real time. Although punch-in/out points can often be manually performed on-the-fly,most sequencers can perform a punch automatically, once the in/out measure numbers have been graphically or numerically entered. The sequence can then be rolled back a few measures and the artist can play along, while the sequencer automatically performs the necessary switching functions (usually with multiple take and full undo capabilities).

In addition to recording a performance in a track-based environment, most sequencers let you enter note values into sequence one note at a time. This feature (known as step time) lets you give the sequencer a basic tempo and note length (i.e., quarter note, sixteenth note, etc.) and then manually enter the notes from a keyboard or other controller. This data entry style is often (but not always) used with fast, high-tech and dance styles, where a real-time performance just isn't possible or accurate enough for the song.

Whether you're recording a track in real time or in step time, it's almost always best to select the proper song tempo before recording a sequence. I bring this up because most sequencers are able to output a click track that can be used as an accurate, audible guide for keeping in time with the song's selected tempo. It's also critical that the tempo be accurate when trying to sync groove loops and rhythms to a sequence via plug-ins or external instruments.

Editing. One of the more important features that a sequencer (or sequenced MIDI track within a DAW) has to offer is its ability to edit tracks or blocks within a track. Of course, these editing functions and capabilities often vary between hardware and software sequencers.

The main track window of a sequencer or MIDI track on a DAW is used to display such track information as the existence of track data, track names, MIDI port assignments for each track, program change assignments, volume controller values, etc.

Depending on the sequencer, the existence of MIDI data on a particular track at a particular measure point (or over a range of measures) is often indicated by the visual display of MIDI data in a piano-roll fashion (showing the general vertical and length placements of the notes as they progress though the musical passage… as shown in Fig. 29-29.

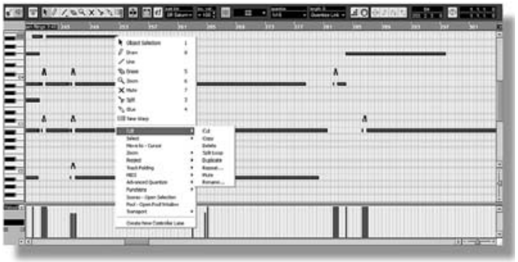


Figure 29-29. The presence of MIDI message data will often appear as a series of highlighted areas within a sequence track or a window. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

By navigating around the various data display and parameter boxes, it's possible to use cut and paste and/or direct edit techniques to vary note, length and controller parameters for almost every facet of a section or musical composition. For example, let's say that we really screwed up a few notes when laying down an otherwise killer bass riff. With MIDI, fixing the problem is totally a no-brainer. Simply highlight each fudged note and drag it to it's proper note location. We can even change the beginning and end points in the process. In addition, tons of other parameters can be changed including velocity, modulation and pitch bend, note and song transposition, quantization, and humanizing (factors that eliminate or introduce human timing errors that are generally present in a live performance), as well as full control over program and continuous controller messages. The list goes on.

Playback. Once a composition is complete, all of the MIDI tracks in a project can be transmitted through the various MIDI ports and channels to plug-ins, instruments, or devices for playback. Since the data exists as encoded real-time control commands, you can listen to the sequence and make changes at any time. For example, you could change instrument settings (by changing or editing patch voices), alter volume and other mix changes, or experiment with such controllers as pitch bend, modulation or aftertouch, even change the tempo and key signature. In short, this medium is infinitely flexible how a performance and/or set of parameters can be created, saved, folded, spindled, and mutilated until you've arrived at the sound and feel that you want.

Another of the greatest beauties of MIDI production is its ability to be altered at any later point in time. For example, let's say that 5 years ago you laid down a killer synth riff in a song that made it onto the charts. A couple of weeks ago a producer came to you in hopes of collaborating on a remix. Of course, technology marches on and your studio has improved over time. First off, even though a lot of the setup parameters have been saved with the original sequence, let's assume that you were smart enough to keep really good setup notes. One big change, however, is that you have a new software synth that has a patch that sounds better than the original patch. Since the remix is to be used in an upcoming film track, MIDI can be used to tweak things up a bit by splitting the riff into two parts: one that contains the lower notes and another the highs. By sending the lows to one patch on the synth and the highs to another, not only have you improved the overall sound, you've filled it out by expanding the soundfield into surround. Without MIDI, you'd have to arrange for a new session and hope that it all goes well, with MIDI, the performance is exactly the same and improvements are made in a no-brainer environment. This is what MIDI's all about—performance, repeatability, easy editing, and cost-effective power!

I now have to take time out to give you a few pointers that will make your life easier when dealing with MIDI production.

1. Remember to set the session to the proper tempo at the beginning of the session. Although tempo can

be changed at a later time, attention to tempo details can help you to avoid later pitfalls.

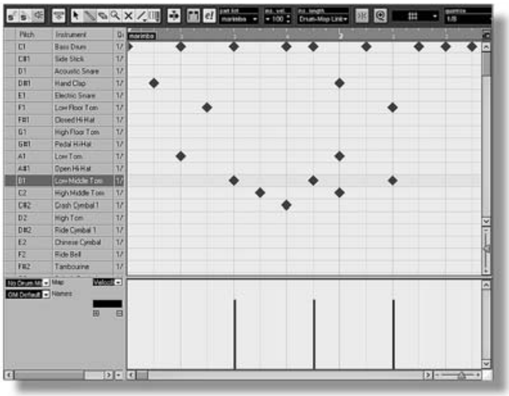
2. Always name your track before you go into record (this goes for both audio and MIDI tracks). Properly naming your tracks (i.e., with its instrument, patch name) is the first step toward good documentation.
3. You can *never* overdocument a session. Keeping good instrument, patch, settings, musician, studio, and other notes might not only come in handy—it can save your butt if you need to revisit the tracks in the future.
4. Never delete a final take MIDI track from a DAW session. Even though you've transferred the instrument to an audio track, it is *always* wise to archive the original MIDI track with session. Trust me, both you and the producer will be glad you did, should any changes need to be made to the track in the future.

29.4.2 Other Software Sequencing Applications

In addition to DAW and sequencing packages that are designed to handle most of the day-to-day production needs of the musician, other types of software tools and applications exist that can help to carry out specialized tasks. A few of these packages include drum pattern editors, algorithmic composition programs, patch editors and music printing programs.

Drum-Pattern Editor/Sequencers. At any one time, there are a handful of companies that have software or hardware devices, that are specifically designed to create and edit, drum patterns. In addition, most of the higher-end DAW audio production systems also include a drum pattern editor that relies on user input and quantization to construct and chain together any number of user-created percussion grooves. More often than not, these editors use a grid pattern that displays drum-related MIDI notes or subpatterns along the vertical axis, while time is represented in metric divisions along the horizontal axis, Fig. 29-30. By clicking on each grid point with a mouse or other input system, individual drum or effect sounds can be built into rhythmic patterns.

Once created, these and other patterns can be linked together to create a partial or complete rhythm section within a song. These editors commonly offer such features as the ability to change MIDI note values (thereby changing drum voices), note length, quantization and humanization, as well as adjustments to note and pattern velocities. Once completed, the sequenced drum track (or chained patterns) can be imported into a sequence, saved, and/or exported.



Pitch	Instrument
C1	Bass Drum
C#1	Side Stick
D1	Acoustic Snare
D#1	Hand Clap
E1	Electric Snare
F1	Low Floor Tom
F#1	Closed Hi-Hat
G1	High Floor Tom
G#1	Pedal Hi-Hat
A1	Low Tom
A#1	Open Hi-Hat
B1	Low-Middle Tom
C2	High Middle Tom
C#2	Crash Cymbal 1
D2	High Tom
D#2	Ride Cymbal 1
E2	Chinese Cymbal
F2	Ride Bell
F#2	Tambourine

Figure 29-30. Steinberg Cubase/Nuendo drum edit window. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

Groove Tools. Getting into the groove of a piece of music often refers to a feeling that's derived from the underlying foundation of the piece: rhythm. With the introduction and maturation of MIDI and digital audio, new and wondrous tools have made their way into the mainstream of music production that can help us to use these technologies to forge, fold, mutilate and create compositions that make direct use of rhythm and other building blocks of music through the use of looping technology.

Of course, the cyclic nature of loops can be— repeat repeat—repetitive in nature, but new toys and techniques in looping have injected the notion of flexibility, real-time control, real-time processing, and mixing to new heights that can be used by an artist as a wondrously expressive tool.

Loop-based audio editors are groove-driven music programs, Figs. 29-31 and 29-32, that are designed to let you drag and drop prerecorded or user-created loops and audio tracks into a graphic multitrack production interface. At their basic level, these programs differ conceptually from their traditional DAW counterpart, in that the pitch- and time-shift architecture is so variable and dynamic that even after the basic rhythmic, percussive and melodic grooves have been created, their tempo, track patterns, pitch, session key, etc. can be quickly and easily changed at any time. With the help of custom, royalty-free loops (available from the manufacturer and/or third-party companies), users can quickly and easily experiment with setting up grooves, backing tracks, and creating a sonic ambience by simply drag-

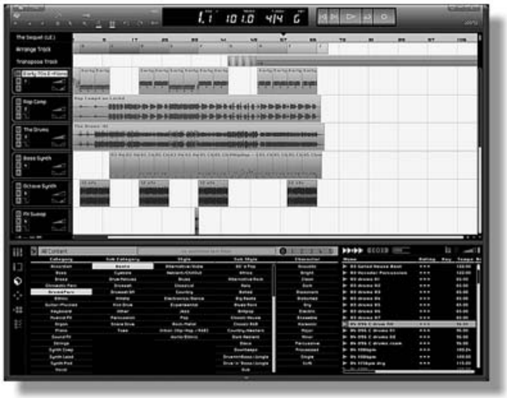
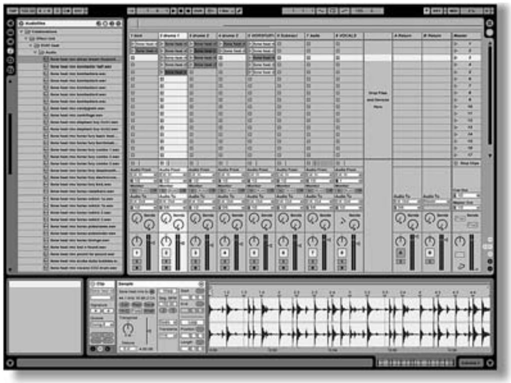


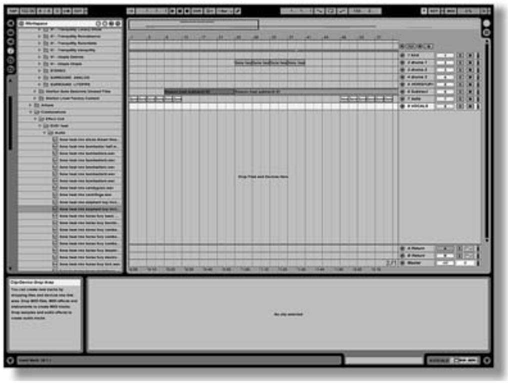
Figure 29-31. Seinberg's Sequel music software. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

ging the loops into the program's main soundfile view where they can be arranged, edited, processed, saved, and exported.

One of the most interesting aspects of a loop-based editor is its ability to match the tempo of a specially programmed loop soundfile to the tempo of the current session. Amazingly enough, this process isn't that difficult to perform, as the program extracts the length, native tempo, and pitch information from the imported file's header and (using various digital time and/or pitch change techniques) adjusts the loop to fit the native time/pitch parameters of the current session. This means that loops of various tempos and musical keys can be automatically adjusted in length and pitch so as to fit in time with previously existing loops. These shifts in time to match a loop to the session's native tempo can actually be performed in a number of ways. For example, using basic DSP techniques to time-stretch and pitch-shift a recorded loop will often work well over a given plus-or-minus percentage range (which is often dependent on the quality of the program algorithms). Beyond this range, the loop will often begin to distort and become jittery. At such extremes, other playback algorithms and beat slice detection techniques can be used to make the loop sound more natural. For example, drums or percussion can be stretched in time by adding additional silence between the various hit points within the loop at precisely calculated intervals. In this way, the pitch will remain the same while the length is altered. Of course, such a loop would sound choppy and broken up when played on its own; however, when buried within a mix, it might work just fine. It's all up to you and the current musical context.



A. Arrangement view.



B. Session view.

Figure 29-32. Ableton live performance audio workstation. Courtesy of Ableton, www.ableton.com.

The software world doesn't actually hold the total patent on looping tools and toys; there are a number of groove keyboards and module boxes that are on the market. These systems, which range widely in sounds, functionality, and price, can offer up a wide range of unique sounds that can be quite useful laying a foundation under your production. In the past, getting a hardware grove tool to sync into a session could be time-consuming, frustrating, and problematic, taking time and tons of manual reading. However, with the advent of powerful time and pitch shift processing within most DAWs, the sounds from these hardware devices can be pulled into a session without too much trouble. For example, a single groove loop (or multiple loops) could be recorded into a DAW (at a bpm that's near to the session's tempo), edited, and then imported into the session, at which time the loop could be easily stretched into time sync, allowing it to be looped to your heart's content. Just remember, necessity is the mother of invention. Patience and creativity are probably your most important tools in the looping process.

If there's a software package that has gripped the hearts and minds of electronic musicians in the 21st century, it would have to be Reason from the folks at Propellerheads, Fig. 29-33. Reason defies specific classification in that it's an overall music production environment that has many facets. For example, it includes a MIDI sequencer, as well as a wide range of software instrument modules that can be played, mixed, and combined in a comprehensive environment that can be controlled from any external keyboard and/or MIDI controller. Reason also includes a large number of signal processors that can be applied to any instrument or instrument group under full and easily controlled automation.

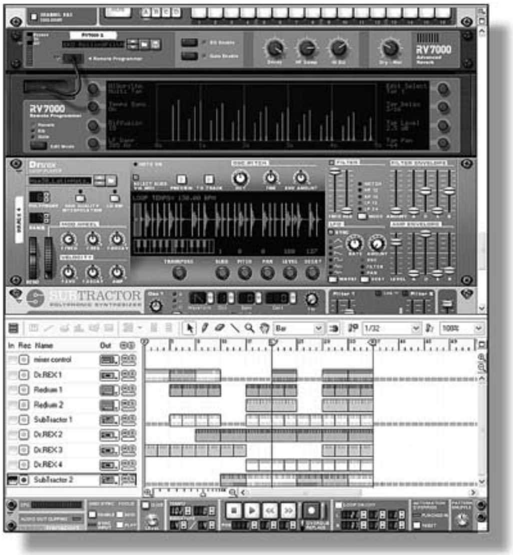


Figure 29-33. Reason music production environment. Courtesy of Propellerheads software, www.propellerheads.se.

In essence, Reason is a combination of modeled representations of vintage analog synthesis gear, mixed with the latest digital synthesis and sampling technology. Combine these with a modular approach to signal and effects processing; add a generous amount of internal and remote mix and controller management (via an external MIDI controller); top this off with a quirky but powerful sequencer; and you have a software package that's powerful enough for top-flight production and convenient enough that you can build tracks from your laptop from your seat in a crowded plane. I know that it sounds like I read this from a sales brochure, but these are the basic facts of this program. When asked to explain Reason to others, I'm often at a loss as the basic structure is so open-ended and flexible that the program can be approached in as many ways as there are people who produce on it. That's not to say that Reason doesn't have a signature sound—it often does. However, it's a tool that can be either used on its own or in combination with other production instruments and tools.

Algorithmic Composition Programs. Algorithmic composition programs are interactive sequencers that directly interface with MIDI controllers or imported files to generate a performance in real time, according to user-programmed computer parameters. In short, once you give it a few basic musical guidelines, it can act as a compositional robot to generate performances or musical parts on its own in order to help you gain new ideas for a song, create an automatic accompaniment, make improvisational exercises, create special performances, or just plain have fun.

This type of sequencer can be programmed to control the performance according to musical key, generated notes, basic order, chords, tempo, velocity, note density, rhythms, accents, etc. Alternatively, an existing standard MIDI file can be imported and further manipulated in real time, according to new parameters that can be varied from a computer keyboard, mouse, or controller. Often such interactive sequencers will accept input from multiple players, allowing it to be performed as a collective jam. Once a composition has been satisfactorily generated, a standard MIDI file can be created and imported into any sequencer.

Patch Editors. The vast majority of MIDI instruments and devices store their internal patch data within RAM memory. Synths, samplers, or other devices contain information on how to configure oscillators, amplifiers, filters, tuning, and other presets in order to create a particular sound timbre or effect. In addition to controlling sound patch parameters, a unit's internal memory can also store such setup information as effects processor settings, keyboard splits, MIDI channel routing, controller assignments, etc.

Although these settings can be manually accessed from the device's panel controls, another (and sometimes more straightforward) way to gain real-time control over the parameters of an instrument or devices is through the use of a patch editor, Fig. 29-34. A patch editor is a software package that's used to provide on-screen controls and graphic windows for emulating and varying an instrument's parameter controls in real time.

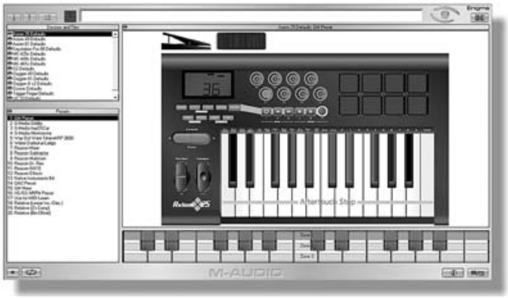


Figure 29-34. M-Audio Enigma Software Librarian and Editor. Courtesy of M-Audio, a division of Avid Technology, Inc., www.m-audio.com.

Direct communication between a patch editor and the device's microprocessor commonly occurs through the use of MIDI SysEx messages. Almost all popular voice and setup editing packages include provisions for receiving and transmitting bulk patch data in this way. This makes it possible to save and organize large numbers of patch-data files, vary setting in real time, and print out patch parameter settings.

In addition to software editing packages, there are also hardware solutions for gaining quick and easy access to device parameters via SysEx. In recent years, MIDI data controllers, Fig. 29-35, have sprung onto the market that can control a wide range of instruments and devices using data faders and soft buttons to vary patch, system, and performance parameters, in real time. In many situations, these controllers can also be used to directly control the volume and mix parameters of a DAW.

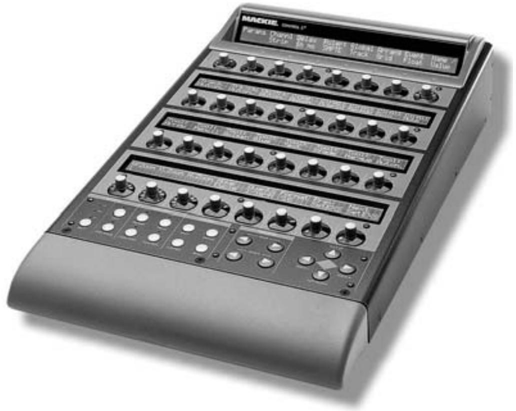


Figure 29-35. Mackie C4 plug-in and virtual instrument controller. Courtesy of Loud Technologies, Inc., www.mackie.com.

Music-Printing Programs. In recent years, the field of transcribing musical scores onto paper has been strongly affected by computer, DAW, and MIDI technology. This process has been enhanced through the use of newer generations of software that make it possible for music notation data to be entered into a computer either manually (by placing the notes onto the screen via keyboard and/or by mouse movements) or via direct MIDI input. Once entered, these notes can be edited in an on-screen environment using a music printing program (or notation app within a DAW) that lets you change and configure a musical score or lead sheet using standard cut-and-paste edit techniques. In addition, most printing programs can play the various instruments in a MIDI system directly from the score. A final and important program feature is their ability to print out hard copies of a score or lead sheets in a wide number of print formats and styles.

These programs or DAW program apps, Fig. 29-36, allow musical data to be entered into a computerized score in a number of manual and automated ways (often with varying degrees of complexity and ease). Although scores can be manually entered, most music-transcription programs will generally accept direct MIDI input, allowing a part to be played directly into a sequence. This can be done in real time (by playing a MIDI instrument or finished sequence into the program), in step time (entering the notes of a score one note at a time from a MIDI controller), or from an existing standard or program-specific MIDI file.

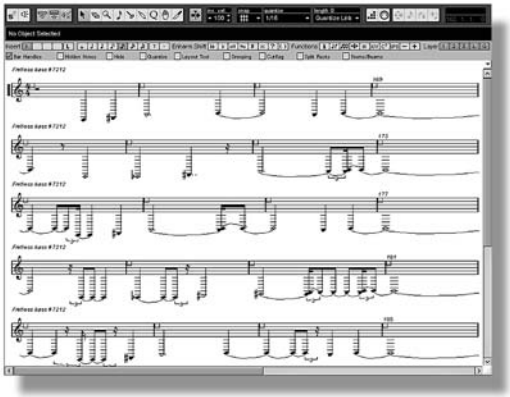


Figure 29-36. Score application within Steinberg's Nuendo DAW software. Courtesy of Steinberg Media Technologies GmbH, a division of Yamaha Corporation, www.steinberg.net.

Another way to enter music into a score is through the use of an optical recognition program. These programs let you place sheet music or a printed score onto a standard flatbed scanner, scan the music into a program and then save the notes and general layout as a NIFF (notation interchange file format) file.

One of the biggest drawbacks to automatically entering a score via MIDI (either as a real-time performance or from a MIDI file) is the fact that music notation is a very interpretive art. "To err is human," and it's commonly this human feel that gives music its full range of expression. It is very difficult, however, for a program to properly interpret these minute yet important imperfections and place the notes into the score exactly as you want them. (For example, it might interpret a held quarter-note as either a dotted quarter-note or one that's tied to a thirty-second note.) Even though these computer algorithms are getting better at interpreting musical data and quantization can be used to tell a computer to round a note value to a specified length, i.e., a score will still often need to be manually edited to correct for misinterpretations.

29.5 Multimedia and the Web

It's no secret that modern-day computers have gotten faster, sleeker, and sexier in their overall design. In addition to its ability to act as a multifunctional production workhorse, one of the crowning achievements of the modern computer is the degree of media and networking integration that has worked its way into our collective consciousness and become known as multimedia.

The combination of working and/or playing with multimedia has found its way into modern computer culture through the use of various hardware and software systems that work in a multitasking environment and combine to bring you a unified experience that seamlessly involves such media types as:

- Text.
- Graphics.
- Video.
- Audio and music.
- Computer animation.
- MIDI.

The obvious reason for integrating and creating these media types is the human desire to create content with the intention of sharing and communicating one's experiences with others. This has been done for centuries in the form of books and more recently by movies and television. In the here and now, the Web has been added to the communications list, in that it has created a vehicle that allows individuals (and corporate entities alike) to communicate a multimedia experience to millions and then allows each individual to manipulate that experience, learn from it, and even respond in an interactive fashion. The Web has indeed unlocked the potential for experiencing multimedia events and information in a way that makes each of us a participant, not just a passive spectator.

One of the unique advantages of MIDI, as it applies to multimedia, is the rich diversity of musical instruments and program styles that can be played back in real time while requiring almost no overhead processing from the computer's CPU. This makes MIDI a perfect candidate for playing back soundtracks from multimedia games or over the Internet. It's interesting to note that MIDI has taken a back seat to digital audio as a serious music playback format for multimedia. Most likely, this is due to several factors, including:

1. A basic misunderstanding of the medium.
2. The fact that producing MIDI content requires a basic knowledge of music.
3. The frequent difficulty of synchronizing digital audio to MIDI in a multimedia environment.
4. The fact that soundcards often include poorly designed FM synthesizers (although most operating systems now include a higher-quality software synth).

Fortunately, an increasing number of software companies have taken up the banner of embedding MIDI within their media projects and have helped push MIDI a bit more into the Web and gaming mainstream. As a result, it's becoming more common for your PC to begin playing back a MIDI score on its own or perhaps in conjunction with a more data-intensive program or game.

29.5.1 Standard MIDI Files

The accepted format for transmitting files or real-time MIDI information in multimedia (or between sequencers from different manufacturers) is the standard MIDI file. This file type (which is stored with a .mid or .smf extension) is used to distribute MIDI data, song, track, time signature, and tempo information to the general masses. Standard MIDI files can support both single and multichannel sequence data and can be loaded into, edited, and then directly saved from almost any sequencer package. When exporting a standard MIDI file, keep in mind that they come in two basic flavors: type 0 and type 1.

- Type 0 is used whenever all of the tracks in a sequence need to be compressed into a single MIDI track. All of the original channel messages still reside within that track; however, the data will have no definitive track assignments. This type might be the best choice when creating a MIDI sequence for the Internet (where the sequencer or MIDI player application might not know or care about dealing with multiple tracks).

- Type 1, on the other hand, will retain its original track information structure and can be imported into another sequencer type with its basic track information and assignments left intact.

29.5.2 *General MIDI*

One of the most interesting aspects of MIDI production is the absolute setup and patch uniqueness of each professional and even semipro project studio. In fact, no two studios will be alike (unless they've been specifically designed to be the same or there's some unlikely coincidence). Each artist will be unique in having his or her own favorite equipment, supporting hardware, favorite way of routing channels and tracks, and assigning patches. The fact that each system setup is unique and personal has placed MIDI at odds with the need for systems compatibility in the world of multimedia. For example, after importing a standard MIDI file over the Net and loading it into a sequencer, you might hear a song that's being played with a totally irrelevant set of sound patches (it might sound interesting, but it won't sound anything like it was originally intended). If the MIDI file is loaded into a new computer, the sequence will again sound completely different, with patches that are so irrelevant that the guitar track might sound like a bunch of machine-gun shots from the planet Gloob.

In order to eliminate (or at best reduce) the basic differences that exist between systems, a patch and settings standard known as General MIDI (GM) was created. In short, GM assigns a specific instrument patch to each of the 128 available program change numbers. Since all electronic instruments that conform to the GM format must use these patch assignments, placing GM program change commands at the header of each track will automatically configure the sequence to play with its originally intended sound. As such, no matter what sequencer is used to play the file back, as long as the receiving instrument conforms to the GM spec the sequence will be heard using its intended instrumentation. Tables 29-3 and 29-4 detail the program numbers and patch names that conform to the GM format (Table 29-3 for percussion and Table 29-4 for nonpercussion instruments). These patches include sounds that imitate synthesizers, ethnic instruments, and/or sound effects that have been derived from early Roland synth patch maps. Although the GM spec states that a synth must respond to all 16 MIDI channels, the first nine channels are reserved for instruments, while GM restricts the percussion track to MIDI channel 10.

Table 29-3. GM percussion instrument patch map (Channel 10)

35. Acoustic Bass Drum
36. Bass Drum 1
37. Side Stick
38. Acoustic Snare
39. Hand Clap
40. Electric Snare
41. Low Floor Tom
42. Closed Hi-Hat
43. High Floor Tom
44. Pedal Hi-Hat
45. Low Tom
46. Open Hi-Hat
47. Low-Mid Tom
48. Hi Mid Tom
49. Crash Cymbal 1
50. High Tom
51. Ride Cymbal 1
52. Chinese Cymbal
53. Ride Bell
54. Tambourine
55. Splash Cymbal
56. Cowbell
57. Crash Cymbal 2
58. Vibraslap
59. Ride Cymbal 2
60. Hi Bongo
61. Low Bongo
62. Mute Hi Conga
63. Open Hi Conga
64. Low Conga
65. High Timbale
66. Low Timbale
67. High Agogo
68. Low Agogo
69. Cabasa
70. Maracas
71. Short Whistle
72. Long Whistle
73. Short Guiro
74. Long Guiro
75. Claves
76. Hi Wood Block
77. Low Wood Block
78. Mute Cuica
79. Open Cuica
80. Mute Triangle
81. Open Triangle

Note: In contrast to Table 29-3, the numbers in Table 29-4 represent the percussion keynote numbers on a MIDI keyboard, not program change numbers.

29.6 MIDI-Based Synchronization

Just as synchronization is routinely used in audio and video production, the wide acceptance of MIDI and digital audio within the various media has created the need for synchronization in project studio and midsized production environments. Devices such as MIDI sequencers, digital audio editors, effects devices, and digital mixing consoles make extensive use of synchronization and time code. However, advances in design have fashioned this technology into one that's much more cost-effective and easy-to-use—all through the use of MIDI. The following sections outline the various forms of synchronization that are often encountered in a MIDI-based production environment.

Simply stated, most current forms of synchronization use the MIDI protocol itself for the transmission of sync messages. These messages are transmitted along with other MIDI data over standard MIDI cables, with no need for additional or special connections.

Table 29-4. GM Non-percussion Instrument Patch Map with Program Change Numbers

1. Acoustic Grand Piano
2. Bright Acoustic Piano
3. Electric Grand Piano
4. Honky-tonk Piano
5. Electric Piano 1
6. Electric Piano 2
7. Harpsichord
8. Clavi
9. Celesta
10. Glockenspiel
10. Music Box
12. Vibraphone
13. Marimba
14. Xylophone
15. Tubular Bells
16. Dulcimer
17. Drawbar Organ
18. Percussive Organ
19. Rock Organ
20. Church Organ
21. Reed Organ
22. Accordion
23. Harmonica
24. Tango Accordion
25. Acoustic Guitar (nylon)
26. Acoustic Guitar (steel)
27. Electric Guitar (jazz)
28. Electric Guitar (clean)
29. Electric Guitar (muted)
30. Overdriven Guitar
31. Distortion Guitar
32. Guitar harmonics
33. Acoustic Bass
34. Electric Bass (finger)
35. Electric Bass (pick)
36. Fretless Bass
37. Slap Bass 1
38. Slap Bass 2
39. Synth Bass 1
40. Synth Bass 2
41. Violin
42. Viola
43. Cello
44. Contrabass
45. Tremolo Strings
46. Pizzicato Strings
47. Orchestral Harp
48. Timpani
49. String Ensemble 1
50. String Ensemble 2
51. SynthStrings 1
52. SynthStrings 2
53. Choir Aahs
54. Voice Oohs
55. Synth Voice
56. Orchestra Hit
57. Trumpet
58. Trombone
59. Tuba
60. Muted Trumpet
61. French Horn
62. Brass Section
63. SynthBrass 1
64. SynthBrass 2
65. Soprano Sax
66. Alto Sax
67. Tenor Sax
68. Baritone Sax
69. Oboe
70. English Horn
71. Bassoon
72. Clarinet
73. Piccolo
74. Flute
75. Recorder
76. Pan Flute
77. Blown Bottle
78. Shakuhachi
79. Whistle
80. Ocarina
81. Lead 1 (square)
82. Lead 2 (sawtooth)
83. Lead 3 (calliope)
84. Lead 4 (chiff)
85. Lead 5 (charang)
86. Lead 6 (voice)
87. Lead 7 (fifths)
88. Lead 8 (bass þ lead)
89. Pad 1 (new age)
90. Pad 2 (warm)
91. Pad 3 (polysynth)
92. Pad 4 (choir)
93. Pad 5 (bowed)
94. Pad 6 (metallic)
95. Pad 7 (halo)
96. Pad 8 (sweep)
97. FX 1 (rain)
98. FX 2 (soundtrack)
99. FX 3 (crystal)
100. FX 4 (atmosphere)
101. FX 5 (brightness)
102. FX 6 (goblins)
103. FX 7 (echoes)
104. FX 8 (sci-fi)
105. Sitar
106. Banjo
107. Shamisen
108. Koto
109. Kalimba
110. Bag pipe
111. Fiddle
112. Shanai
113. Tinkle Bell
114. Agogo
115. Steel Drums
116. Woodblock
117. Taiko Drum
118. Melodic Tom
119. Synth Drum
120. Reverse Cymbal
121. Guitar Fret Noise
122. Breath Noise
123. Seashore
124. Bird Tweet
125. Telephone Ring
126. Helicopter
127. Applause
128. Gunshot

29.6.1 MIDI Real-Time Messages

Although no time code–based reference is implemented, it's important to know that MIDI has a built-in (and often transparent) protocol for synchronizing all of the tempo and timing elements of each MIDI device in a system to a master clock. This protocol operates by transmitting real-time messages to the various instruments and devices throughout the system. Although these relationships are usually automatically defined within a system setup, one MIDI device must be designated as the master device in order to provide the timing information to which all other slaved devices are locked. MIDI real-time messages consist of four basic types that are each 1 byte in length:

- **Timing clock—**A clock timing that's transmitted to all devices in the MIDI system at a rate of 24 pulses per quarter note (ppq). This method is used to improve the system's timing resolution and simplify timing when working in nonstandard meters (e.g., $^{3}/_{8}$, $^{5}/_{16}$, $^{5}/_{32}$).
- **Start**—Upon receipt of a timing clock message, the start command instructs all connected devices to begin playing from the beginning of their internal sequences. Should a program be in midsequence, the

start command repositions the sequence back to its beginning, at which point it begins to play.

- **Stop**—Upon the transmission of a MIDI stop command, all devices in the system stop at their current positions and wait for a message to follow.
- **Continue**—Following the receipt of a MIDI stop command, a MIDI continue message instructs all instruments and devices to resume playing from the precise point at which the sequence was stopped. Certain older MIDI devices (most notably drum machines) aren't capable of sending or responding to continue commands. In such a case, the user must either restart the sequence from its beginning or manually position the device to the correct measure.

29.6.2 Song Position Pointer

In addition to MIDI real-time messages, the Song Position Pointer (SPP) is a MIDI system common message that isn't commonly used in current-day production. Essentially, SPP keeps track of the current position in the song by noting how many measures have passed since the beginning of a sequence. Each pointer is expressed as multiples of six timing-clock messages and is equal to the value of a 16th note.

The song position pointer can synchronize a compatible sequencer or drum machine to an external source from any position within a song containing 1024 or fewer measures. Thus, when using SPP, it is possible for a sequencer to chase and lock to a multitrack tape from any measure point in a song.

Using such a MIDI/tape setup, a specialized sync tone is transmitted that encodes the sequencer's SPP messages and timing data directly onto tape as a modulated signal. Unlike SMPTE time code, the encoding method wasn't standardized between manufacturers. This lack of standardization prevents SPP data written by one device from being decoded by another device that uses an incompatible proprietary sync format.

Unlike SMPTE, where tempos can be easily varied by inserting a tempo change at a specific SMPTE time, once the SPP control track is committed to tape, the tape and sequence are locked into this predetermined tempo or tempo change map. SPP messages are usually transmitted only while the MIDI system is in the stop mode, in advance of other timing and MIDI continue messages. This is due to the relatively short time period that's needed to locate the slaved device to the correct measure position.

29.6.3 MIDI Time Code

MIDI time code (MTC) was developed to allow electronic musicians, project studios, video facilities, and virtually all other production environments to cost-effectively and easily translate time code into time-stamped messages that can be transmitted via MIDI. Created by Chris Meyer and Evan Brooks, MTC enables SMPTE-based time code to be distributed throughout the MIDI chain to devices or instruments that are capable of synchronizing to and executing MTC commands. MTC is an extension of MIDI 1.0, which makes use of existing message types that were either previously undefined or were being used for other nonconflicting purposes. Since most modern recording devices include MIDI in their design, there's often no need for external hardware when making direct connections. Simply chain the MIDI cables from the master to the appropriate slaves within the system (via physical cables, USB, or virtual internal routing). Although MTC uses a reasonably small percentage of MIDI's available bandwidth (about 7.68% at 30 fr/s), it's customary (but not necessary) to separate these lines from those that are communicating performance data when using MIDI cables. As with conventional SMPTE, only one master can exist within an MTC system, while any number of slaves can be assigned to follow, locate, and chase to the master's speed and position. Because MTC is easy to use and is often included free in many system and program designs, this technology has grown to become the most common and most straightforward way to lock together such devices as DAWs, modular digital multitracks, and MIDI sequencers, as well as analog and videotape machines (by using a MIDI interface that includes a SMPTE-to-MTC converter).

The MTC format can be divided into two parts:

- Time code.
- MIDI cueing.

The time code capabilities of MTC are relatively straightforward and allow devices to be synchronously locked or triggered to SMPTE time code. MIDI cueing is a format that informs a MIDI device of an upcoming event that's to be performed at a specific time (such as load, play, stop, punch in/out, reset). This protocol envisions the use of intelligent MIDI devices that can prepare for a specific event in advance and then execute the command on cue.

MTC is made up of three message types: quarter-frame messages, full messages, and MIDI cueing messages.

- **Quarter-frame messages**—These are transmitted only while the system is running in real or variable speed time, in either forward or reverse direction. True to its name, four quarter-frame messages are generated for each time code frame. Since 8 quarter-frame messages are required to encode a full SMPTE address (in hours, minutes, seconds, and frames—00:00:00:00), the complete SMPTE address time is updated once every two frames. In other words, at 30 fps, 120 quarter-frame messages would be transmitted per second, while the full time code address would be updated 15 times in the same period. Each quarter frame message contains 2 bytes. The first byte is F1, the quarter-frame common header, while the second byte contains a nibble (four hits) that represents the message number (0 through 7) and a nibble for encoding the time field digit.
- **Full messages**—Quarter-frame messages are not sent in the fast-forward, rewind, or locate modes, as this would unnecessarily clog a MIDI data line. When the system is in any of these shuttle modes, a full message is used to encode a complete time code address within a single message. After a fast shuttle mode is entered, the system generates a full message and then places itself in a pause mode until the time-encoded slaves have located to the correct position. Once playback has resumed, MTC will again begin sending quarter-frame messages.
- **MIDI cueing messages**—MIDI cueing messages are designed to address individual devices or programs within a system. These 13 bit messages can be used to compile a cue or edit decision list, which in turn instructs one or more devices to play, punch in, load, stop, and so on at a specific time. Each instruction within a cueing message contains a unique number, time, name, type, and space for additional information. At the present time, only a small percentage of the possible 128 cueing event types has been defined.

SMPTE/MTC Conversion. Although MTC is commonly implemented within a software or hardware system itself (that's the functional and economic beauty of it), whenever a hardware device that doesn't talk MTC (but only a flavor of the SMPTE protocol), a SMPTE-to-MIDI converter must be used, Fig. 29-37. These conversion systems are available as stand-alone devices or as an integrated part of a multiport MIDI interface/patch bay/synchronizer system. Certain analog and digital multitrack systems include a built-in MTC port within their design, meaning that the machine can be synchronized to a DAW/sequencing system without the need for any additional hardware, beyond a MIDI interface.

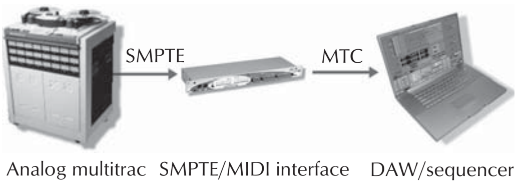


Figure 29-37. SMPTE time code can be easily converted t MTC (and vice versa) for distribution throughout a production system.

This page intentionally left blank

Chapter 30

Optical Disc Formats for Audio Reproduction and Recording

by Ken Pohlmann

30.1 Introduction 1133
30.2 CD Specifications 1133
  30.2.1 Pit Track 1133
  30.2.2 Data Encoding 1134
  30.2.3 Subcode 1135
  30.2.4 EFM Encoding 1137
30.3 CD Player Design 1137
  30.3.1 Optical Pickup 1137
  30.3.2 Data Decoding 1140
  30.3.3 Signal Reconstruction 1140
30.4 Other CD Formats 1141
  30.4.1 CD-ROM 1141
  30.4.2 CD-R 1142
  30.4.3 CD-RW 1143
30.5 Super Audio CD 1144
  30.5.1 SACD Specifications 1144
  30.5.2 Direct Stream Digital Coding 1144
30.6 DVD Disc Format 1145
  30.6.1 DVD Disc Manufacturing 1146
  30.6.2 DVD File Format and Coding 1146
30.7 DVD-Video 1147
  30.7.1 Audio Contents 1148
30.8 DVD-Audio 1148
  30.8.1 File Organization 1149
  30.8.2 Contents and Features 1149
  30.8.3 Meridian Lossless Packing (MLP) 1151
30.9 Other DVD Formats 1151
30.10 HD DVD Format 1152
30.11 Blu-ray Disc Format 1153
References 1154
Bibliography 1154

This page intentionally left blank

30.1 Introduction

The digital storage of audio signals presents a technical challenge. A 60 minute stereo musical selection, with a sampling rate of 44.1 kHz and 16 bit pulse code modulation, generates over 5 billion bits. To store this data successfully, error correction, synchronization, and modulation may push the total required capacity to over 15 billion bits. Recordings with a higher sampling frequency, longer word-length, and additional channels require correspondingly greater storage capacity. In addition, commercial music storage media must provide random access, small size, convenience, durability, low cost, and ease of replication. Still other applications call for write-once or recordable/erasable storage. Clearly, digital audio's storage requirements are formidable.

The CD was the first format able to meet these demands. A CD can hold over an hour of high-fidelity music on a robust and economically manufactured disc. CD player hardware specifications have evolved to surpass most listeners' ability to audibly distinguish between players. In short, the format is well suited to the storage demands of stereo music. In addition, the CD format is suitable for many extended applications. As a result, a number of alternative CD formats were developed. A CD-ROM disc may hold several hours of music, along with video and text information. Write-once and recordable/erasable formats (CD-R and CD-RW) are widely used in both professional and consumer applications.

The desire for higher performance specifications and multichannel sound stimulated development of the super audio CD (SACD) format; it uses direct stream digital coding in place of PCM coding to store either stereo or multichannel audio signals on a multilayer disc. The need for increased storage capacity, particularly for the storage of high-quality digital video, encouraged development of the DVD format. A DVD disc may store from 4.7 to 17 Gbytes of data, using one or multiple data layers. As with the CD, DVD comes in many guises. The DVD-Video format is used to store motion pictures, DVD-Audio is used for high-quality stereo and multichannel music, DVD-ROM is used for computer applications, and a variety of DVD formats have been devised for recording applications. The HD DVD and Blu-ray disc formats use shorter wavelength lasers and higher resolution optics to dramatically increase storage density, allowing storage of high-definition video and audio. Disc formats such as these will further extend the opportunities of optical disc storage for professional and consumer applications.

30.2 CD Specifications

The compact disc digital audio (CD-DA) format is sometimes known as the Red Book standard and is codified in the ISO/IEC 908 standard. The diameter of a CD is 120 millimeters (mm) (4.7 in), its center hole diameter is 15 mm (0.59 in), and its thickness is 1.2 mm (0.047 in). The innermost diameter does not hold data; it provides a clamping area for the player to secure the disc to the spindle motor shaft. Data is recorded on a 35.5 mm (1.4 in) wide area. A lead-in area occupies the innermost data radius, and a lead-out area occupies the outermost radius; they contain nonaudio data used to control the player's operation.

A transparent polycarbonate plastic substrate forms most of a disc's 1.2 mm thickness, as shown in Fig. 30-1. Data is physically contained in pits that are impressed on the top surface of the substrate. The pit surface is covered with a very thin 50 nm to 100 nm (nanometer) metal (e.g., aluminum or gold) layer and another thin 10 µm to 30 µm (micrometer) protective plastic layer, with the 5 µm identifying label printed on top. A laser beam is used to read the data. It is applied from below, passes through the transparent substrate, reflects from the metallized pit surface, and passes back through the substrate. The laser beam is focused on the metallized surface embedded inside the disc. Since data on a disc is read by a light beam, playing a CD does not cause wear to the data surface, or pickup.

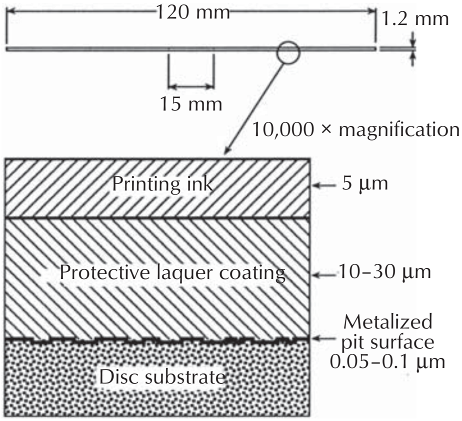


Figure 30-1. CD construction showing substrate, metallized surface, protective layer, and label.

30.2.1 Pit Track

Data is arranged as a pit track in a continuous spiral running from the inner circumference to the outer. A pit is about 0.6 µm wide. A photograph of a pit surface, taken with a scanning electron microscope, is shown in Fig.30-2. The track pitch, the distance between successive tracks, is 1.6 µm; the track pitch acts as a diffraction grating, producing a rainbow of colors. There is a maximum of 20,188 revolutions across the disc's standard data surface width of 35.5 mm.

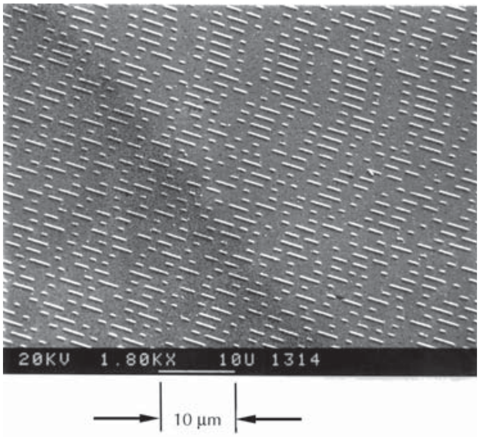


Figure 30-2. Scanning electron microscope photograph of the CD data surface. Courtesy University of Miami.

The linear dimensions of a track are the same at the beginning of a spiral as at the end. This means that a CD rotates with a constant linear velocity (CLV), a condition in which a uniform relative velocity is maintained between the data spiral and the pickup. To accomplish this, the rotation speed of a disc varies depending on the radial position of the pickup. Because each outer track revolution contains more pits than each inner track revolution, the disc must be slowed as it plays outward to maintain a constant rate of data. In particular, the disc rotates at a speed of about 500 rpm when the pickup is reading the inner circumference, and as the pickup moves outward, the rotational speed gradually decreases to about 200 rpm. A constant linear velocity is maintained through a CLV servo system; the player reads frame synchronization from the stored data and varies the disc speed to maintain a constant data rate. The CD standard permits a maximum of 74 minutes, 33 seconds of audio playing time on a disc. However, by reducing parameters such as track pitch and linear velocity, it is possible to manufacture discs with over 80 minutes of music.

The fact that the disc data surface is physically separated from the reading side of the substrate provides a significant asset. Damage and dust on the outer surface do not lie in the focal plane of the reading laser beam and hence their effect is minimized.

The polycarbonate substrate has refractive index of 1.55; the velocity of light slows from 3×10^5 kilometers/second (km/s), to 1.9×10^5 km/s. Because of the bending from the refractive index and thickness of the substrate, and the numerical aperture (NA) of 0.45 of the laser pickup's lens, the diameter of the laser spot is reduced from approximately 800 µm on the disc surface to approximately 1 µm at the pit surface. The laser beam is thus focused to a point larger than a pit width.

The reflective data pit surface, known as land, causes almost 90% of the laser light to be reflected back into the optical pickup. When viewed from the laser's underside perspective, the pits appear as bumps. The height of each bump is between 0.11 and 0.13 µm (110 and 130 nm.) This dimension is slightly smaller than the laser beam's wavelength in air of 780 nm (some players use 790 nm). Inside the polycarbonate substrate, the laser's wavelength is about 500 m. The height of the bumps is thus approximately one-quarter of the laser's wavelength in the substrate.

There is a phase difference between the part of the beam reflected from the bump, and the part reflected from the surrounding land. The phase difference causes destructive interference in the reflected beam. In theory, when the beam strikes an area between pits virtually all of its light is reflected, and when it strikes a pit virtually all of the light returning to the pickup is canceled, hence virtually none is reflected. In practice, the laser spot is larger than required for complete cancellation between pit and land reflections, and pits are made slightly shallower than a quarter wavelength; this yields a better tracking signal, among other things. Typically the presence of a bump reduces reflective power by about 25%. In any case, the data surface varies the intensity of the reflected laser beam. Thus the data physically encoded on the disc can be recovered by the laser, and converted to an electrical signal using a photodiode.

30.2.2 Data Encoding

The audio program played from a CD is the culmination of a data transformation that takes place during master encoding and that undergoes decoding each time the disc is played. Various media are used to hold master recordings. Originally, many CDs were mastered from data recorded on ¾ inch U-matic videotape cassettes using a digital audio processor. In many cases, Exabyte 8 mm data tapes are used to hold the master recording. For audio mastering, the DDP (Disk Description Protocol) file format may be used to hold Red Book and PQ subcode data. Both DDP 1.0 and DDP 2.0 are used; the 2.0 specification writes the TOC to the end of the tape. It is generally recommended to supply a replication plant with an Exabyte tape with DDP files (including PQ and ISRC data). With Exabyte tapes, glass masters may be created at faster than real time speeds. In some cases, audio data is written to a master CD-ROM (CD read-only memory) disc as 24-bit WAV or AIFF files. DAT tapes and CD-R discs can be used as masters, but their relatively higher error rates and susceptibility to damage make them nonideal. An analog tape can also be used as the master. Digital recordings made at a different sampling rate must be passed through a sample rate converter.

CD encoding is the process of placing audio data in a format suitable for storage on the disc. A frame structure provides a means to distinguish the data types. The information contained in a CD frame (prior to modulation) contains a 27-bit sync word, 8-bit subcode, 192 data bits, and 64 parity bits.

Encoding begins with the audio data. Six 32-bit PCM audio sampling periods (alternating from 16-bit left and right channels) are grouped in a frame, left channel preceding right. Each 32-bit sampling period is divided to yield four 8-bit audio symbols. Subsequent signal processing prepares the audio data for storage on the disc surface. In particular, error correction encoding must be accomplished.

The raw error rate from a CD is around 10^{-5} to 10^{-6}, or about one error for every 0.1 to 1.0 million channel (stored) bits. This is impressive storage capability, but considering that a disc outputs 4.3218 million channel bits per second, the need for error correction is obvious. With error correction, 220 errors per second can be completed corrected; interleaving distributes errors, and parity corrects them.

The Cross Interleave Reed-Solomon Code (CIRC) algorithm is used for error correction in the CD system. The CIRC algorithm uses two correction codes for correcting capability, and three interleaving stages to encode data before it is placed on a disc and to decode the data during playback. Because of cross interleaving, the separation of two error correction codes by an interleaving stage, one Reed-Solomon code can check the validity of the other code. The Reed-Solomon code used in CIRC is well suited for the CD system because its decoding requirements are relatively simple. The complete CIRC encoding scheme is shown in Fig. 30-3. With this encoding algorithm, data (twenty-four 8-bit symbols) from the audio signal are cross-interleaved, and two encoding stages generate 8-bits of parity.

30.2.3 Subcode

Following CIRC encoding, an 8-bit CD subcode symbol is added to each frame. The eight subcode bits are designated as P, Q, R, S, T, U, V, and W. Only the P or Q bits are required in the audio format. The CD player collects subcode symbols from 98 consecutive frames to form a subcode block, with eight 98-bit words. Thus the eight subcode bits (P through W) are used as eight different channels with each CD frame containing 1 P bit, 1 Q bit, etc. A subcode block is complete with a synchronization word, instruction and data, commands, and parity. The start of each subcode block is denoted by sync patterns in the first symbol positions of two successive blocks.

The P channel contains a flag bit originally designed for use by simple players to access disc information. In practice, players ignore the P bit and use information in the more comprehensive Q channel. The Q subcode channel is vital for reading audio data on the disc. The Q channel contains four kinds of information: control, address, Q data, and cyclic redundancy check code (CRCC). Each subcode block contains 72-bits of Q data and 16-bits for CRCC, used for error detection on the control, address, and Q data information. The control information flag bits handle several player functions:

1. The number of audio channels (two or four) is indicated; this distinguishes between a two- and four-channel CD recording (the latter not implemented).
2. Preemphasis (on/off) is indicated; a CD track may be encoded with preemphasis, a noise suppression method (this is rarely employed).
3. Digital copy prohibited (yes/no) is indicated.
4. Audio or data content is indicated.

The address information consists of four bits designating the three modes for the Q data bits. Primarily, Mode 1 contains the number and start times of tracks, Mode 2 contains a catalog number, and Mode 3 contains other product codes. Mode 1 stores information in the disc lead-in area, program area, and lead-out area; the data format in the lead-in area differs from that in the other areas. Mode 1 lead-in information is contained in the CD table of contents (TOC). The TOC stores data indicating the number of music selections (up to 99) as a track number and the starting points of the tracks in disc running time. The TOC is read during disc initialization, before the disc begins playing audio data.

In the program and lead-out areas, Mode 1 contains track numbers, indices (subdivision numbers) within a track, time within a track, and absolute time. A time count is set to zero at the beginning of each track and increases to the end of the track. At the beginning of a

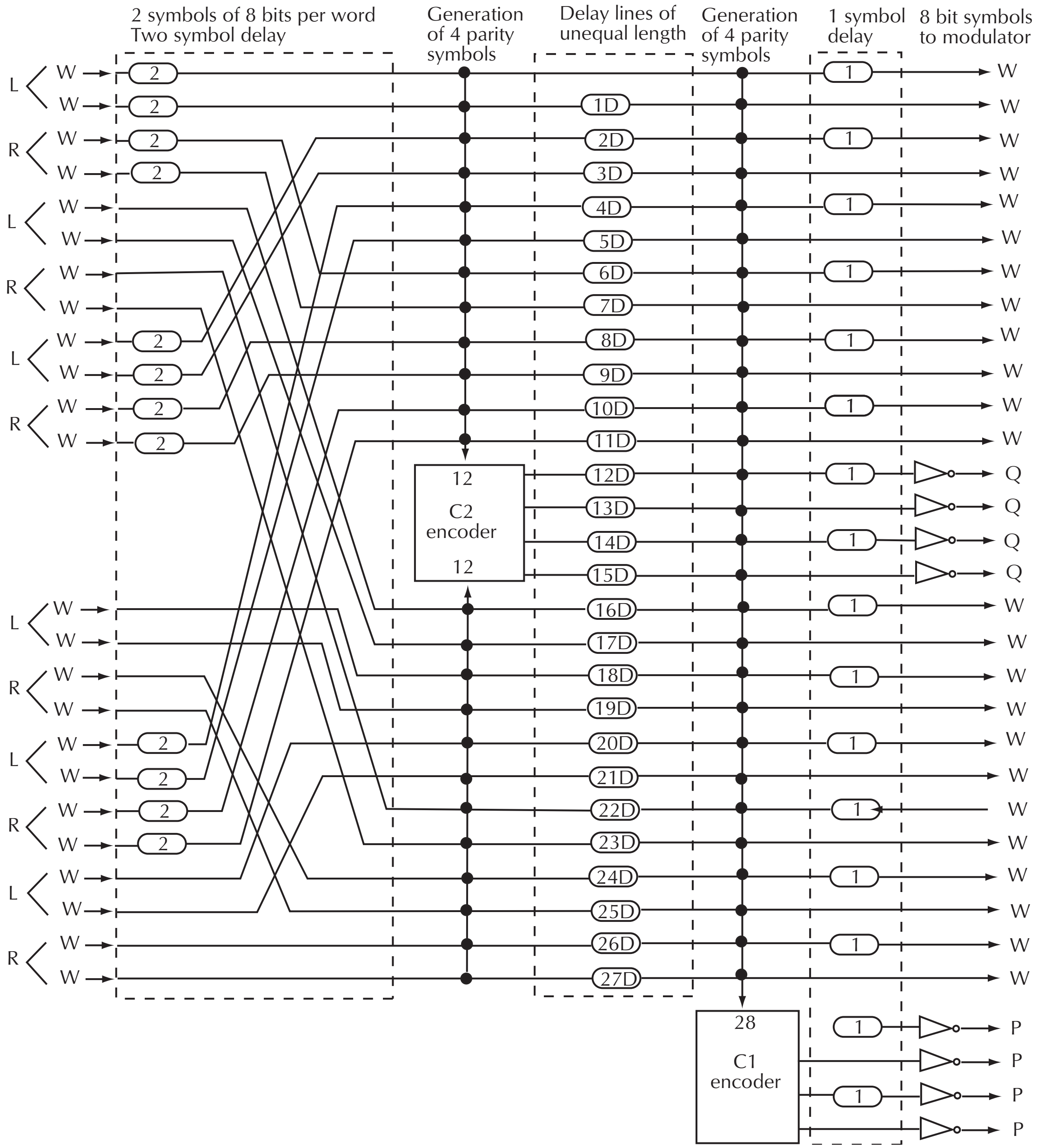

2 symbols of 8 bits per word
Two symbol delay
Generation of 4 parity symbols
Delay lines of unequal length
Generation of 4 parity symbols
1 symbol delay
8 bit symbols to modulator
L
R
W
2
1D
2D
3D
4D
5D
6D
7D
8D
9D
10D
11D
12D
13D
14D
15D
16D
17D
18D
19D
20D
21D
22D
23D
24D
25D
26D
27D
12
C2 encoder
12
28
C1 encoder
1
Q
P

Figure 30-3. CIRC encoding algorithm.

pause, a time count decreases ending with zero at the end of the pause. The absolute time is set to zero at the beginning of the program area and increases to the start of the lead-out area. Time and absolute time are expressed in minutes, seconds, and frames (75 frames per second). Modes 2 and 3 are optional in the subcode.

The other six channels (R, S, T, U, V, and W,) which account for about 20 megabytes of 8-bit storage, are available for other data storage. In some discs, this capacity is used to hold CD-Text, a feature that was appended to the original Red Book specification. With CD-Text, album title, song titles, artist names, and other text information are coded prior to manufacture. Compatible players can read and display CD Text information and also search for particular album titles in a changer. CD-Text supports color bitmaps and JPEG pictures but is generally used to display simple text information.

30.2.4 EFM Encoding

After the audio, parity, and subcode data is assembled, the bit stream is modulated using EFM (eight-to-fourteen modulation). Blocks of 8 data bits are translated into blocks of 14 channel bits, assigning an arbitrary and unambiguous word of 14-bits to each 8-bit word. By choosing select 14-bit words with a low number (and known rate) of 1/0 transitions, greater data density can be achieved. It would be inefficient to store the 8-bit symbols directly on the disc; the large number of 1/0 transitions would demand many pits. In addition, 8-bit symbols have many similar patterns. With 14-bit words, more unique patterns can be selected. EFM thus expedites error correction.

Blocks of 14-bits are linked by three merging bits; two merging bits (always 0s) are required to prevent the possibility of successive 1s between serial words (a violation of the EFM coding scheme). The additional merging bit (either a 1 or a 0, depending on the preceding and succeeding patterns) is added to each code pattern to aid in clock synchronization and to suppress the signal's low-frequency component. The latter is accomplished by selecting merging bits that maintain the signal's average digital sum value at zero. The ratio of bits before and after modulation is 8:17. During demodulation, only 14-bits will be processed, the 3 merging bits are discarded.

The 8 data bits require 2^8 or 256 different code patterns. However the 14-bit channel word can offer 16,384 combinations. To achieve pits of controlled length, only those combinations are selected in which more than two but less than ten 0s appear continuously. In addition, unique patterns are sought. Only 267 combinations satisfy these criteria. Because only 256 patterns are needed, 11 of the 267 patterns are discarded (two of them are used for subcode synchronization words).

The resultant channel stream produces pits and lands that comprise at least two (3T) but no more than ten (11T) successive 0s in length. It is the combination of these varying dimensions that physically encodes the data. The selection of EFM bit patterns defines the physical relationship of the pit dimensions. The pits and intervening reflective land on the CD surface do not directly designate 1s and 0s. Rather, each pit edge whether leading or trailing, is a 1 and all increments in between, whether inside or outside a pit, are 0s, as shown in Fig. 30-4.

With EFM there are more bits to accommodate, but with modulation the highest frequency in the output signal is decreased. Therefore a lower track velocity can be utilized and longer playing time is achieved. This is an efficient encoding method because the number of bits transmitted divided by the number of transitions needed on the medium to convey them is high.

30.3 CD Player Design

A CD player hardware architecture may be considered as five functional elements each working in concert with the other: Optical readout, servo system, spindle motor, control and display, and decoding circuits. The data path directs the modulated light from the pickup through a series of processing circuits, ultimately yielding a stereo analog signal. The data path typically consists of elements such as data separator, deinterleaving RAM, error detection, correction and concealment circuits, oversampling filters, D/A converters, and analog output filters. The servo, control, and display system must direct mechanical operation of the disc, including spindle drive, auto-tracking, and auto-focusing, and handle user interface with the player's controls and displays. A block diagram of the data path is shown in Fig. 30-5.

30.3.1 Optical Pickup

The CD optical pickup must focus, track, and read the data spiral. The entire lens assembly, a combination of the laser source and the reader, must be small enough to

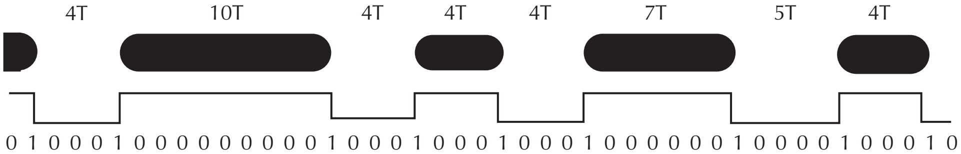


Figure 30-4. Pit/land edges represent logical 1 data.

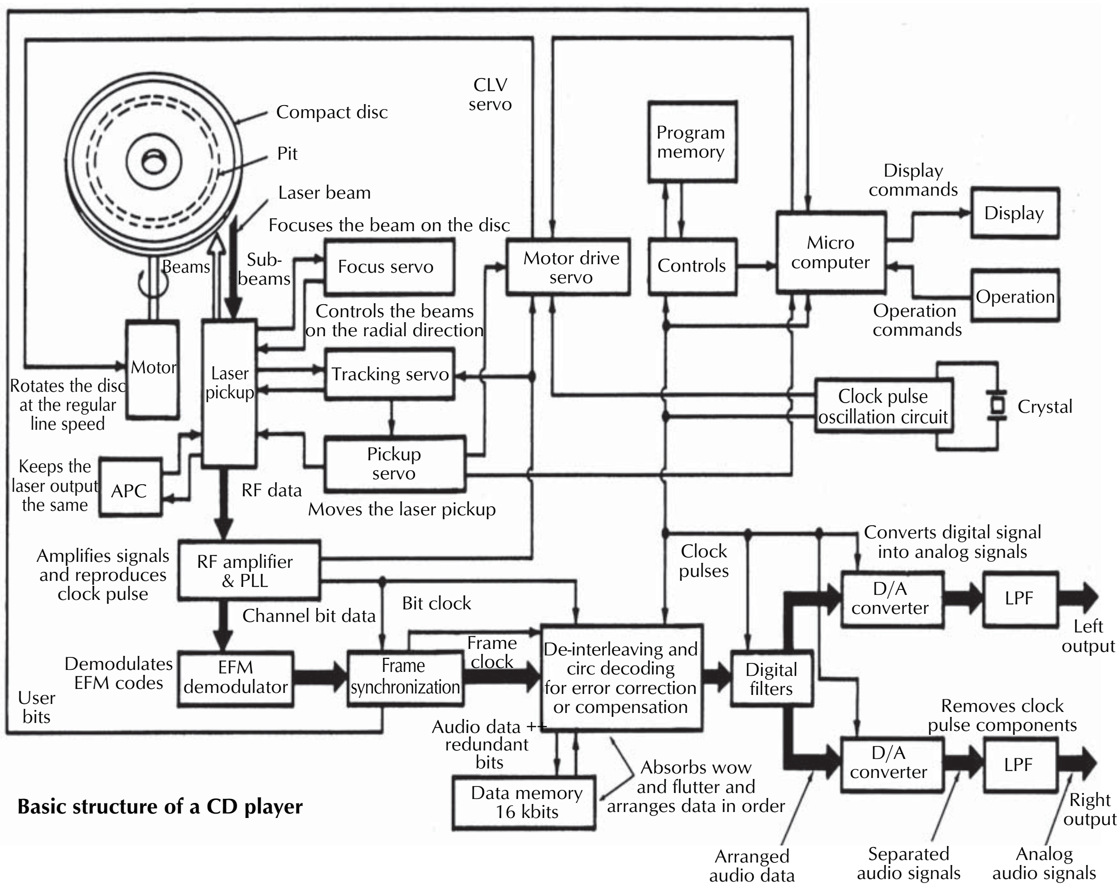

CLV servo
Compact disc
Pit
Laser beam
Focuses the beam on the disc
Program memory
Display commands
Display
Micro computer
Motor drive servo
Controls
Operation
Operation commands
Beams
Sub-beams
Focus servo
Controls the beams on the radial direction
Motor
Laser pickup
Tracking servo
Rotates the disc at the regular line speed
Clock pulse oscillation circuit
Crystal
Pickup servo
Keeps the laser output the same
APC
RF data
Moves the laser pickup
Converts digital signal into analog signals
Amplifies signals and reproduces clock pulse
RF amplifier & PLL
Clock pulses
D/A converter
LPF
Left output
Bit clock
Channel bit data
Frame clock
Demodulates EFM codes
EFM demodulator
Frame synchronization
De-interleaving and circ decoding for error correction or compensation
Digital filters
User bits
Audio data + redundant bits
Removes clock pulse components
D/A converter
LPF
Absorbs wow and flutter and arranges data in order
Data memory 16 kbits
Right output
Basic structure of a CD player
Arranged audio data
Separated audio signals
Analog audio signals

Figure 30-5. CD player block diagram showing optical processing and output signal processing.

glide laterally beneath the disc, moving in response to tracking information and user access demands. Furthermore, the pickup must maintain focusing and tracking even under adverse playing conditions such as a dirty disc or impact and vibration.

To achieve sharp focus on the data surface and intensity modulation, a laser is used as the light source. CD pickups use an AlGaAs semiconductor laser irradiating a coherent-phase laser beam with a 780 nm wavelength (some manufacturers use 790 nm).

CD players can employ either single-beam or three-beam pickups; three-beam designs are more prevalent. A three-beam pickup uses a center beam for reading data and focusing, and two secondary beams for tracking. The design of a three-beam pickup is shown in Fig. 30-6. To generate additional beams, the laser light passes through a diffraction grating, a screen with slits spaced only a few laser wavelengths apart. As the beam passes through the grating, the light diffracts; when the resulting collection is again focused, it will appear as a single bright centered beam with a series of successively less intense beams on either side. Three beams from this diffraction pattern usefully strike the disc. As discussed, when a laser spot strikes land, the smooth interval between two pits, the light is almost totally reflected; when it strikes a pit (seen as a bump by the laser), destructive interference and diffraction causes less light to be reflected into the pickup. The intensity-modulated light is collected by the objective lens and passes through the reading portion of the pickup.

In many three-beam designs, the property of astigmatism is used to achieve auto-focusing. A cylindrical lens is used to detect an out-of-focus condition. As the distance between the objective lens and disc reflective surface varies, the focal point of the optical system also changes, and the image projected by the cylindrical lens changes its shape, as shown in Fig. 30-7. That change in the image on a four-quadrant photodiode generates the focus correction signal. For example, if the disc were too near to the pickup's objective lens, the focal length

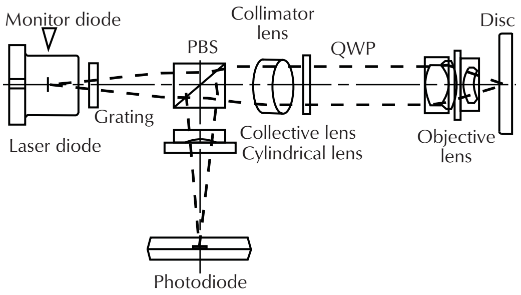


Figure 30-6. Three-beam optical pickup showing diffraction grating, objective lens, and photodiode.

would be shortened and astigmatism from the cylindrical lens would cause the reflected laser spot to be flattened and rotated to one side. This would cause more light to fall on two (opposite) pairs of photodiodes than on the other pairs. This generates a voltage interpreted by the servo system as a command to pull the lens down from the disc. This provides the correct focal path length where astigmatism would not affect the beam. Hence, it would have a round shape, and an equal amount of light would fall on each part of the four-quadrant photodiode, providing a neutral signal to the servo system. When the disc is too far from the lens, the laser spot rotates in the opposite direction, generating a voltage that pushes the lens upward. In practice, the process in this servo loop is a dynamic one, with the objective lens moving in constant accord with disc deviations to provide a correct focal path length.

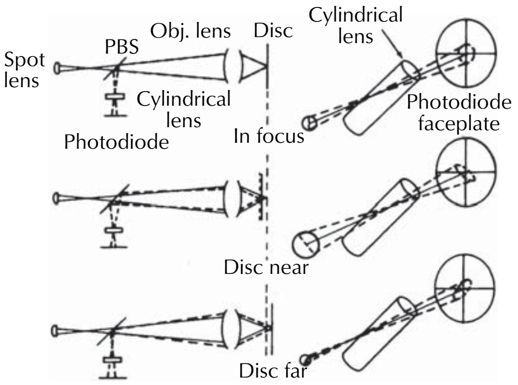


Figure 30-7. Astigmatism used in auto-focusing.

In three-beam pickups, the two secondary beams are used for auto-tracking. The central beam spot covers the pit track while the two tracking beams are aligned above and below and to either side of the center beam. When the beam is tracking the disc properly, part of each tracking beam is aligned on the pit edge; the other part covers the mirrored land between pit tracks. The main beam strikes a four-quadrant photodiode, and the two tracking beams strike two separate photodiodes mounted on either side of the main photodiode.

If the three spots drift to either side of the pit track, the amount of light reflected from the tracking beams varies. There is less average light intensity reflected by the tracking beam that encounters more pit area and greater reflected light intensity from the tracking beam that encounters less pit area. The relative output voltages from the two tracking photodiodes thus form a tracking correction signal, as shown in Fig. 30-8. Operating similarly to the signal used in the auto-focus servo loop, this tracking signal forms a control voltage for the auto-tracking servo mechanism. For example, when the pickup's objective lens drifts to the right of the pit track, the right tracking beam encounters more reflective land and its reflected intensity is greater. When this brighter spot strikes the right tracking photodiode, a voltage greater than that on the left photodiode is generated. This voltage shift causes the servo system to move the pickup to the left, toward the pit track center. Likewise, the opposite occurs when the pickup drifts to the left. In this dynamic process the servo system continually moves the pickup to compensate for track deviations.

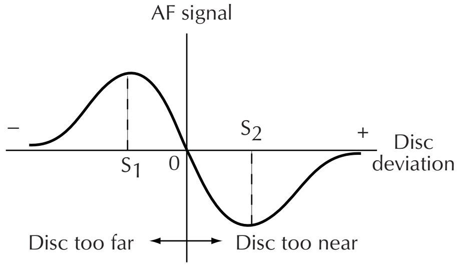


Figure 30-8. Auto-tracking correction signal.

In addition to auto-focus and auto-tracking, a CD pickup uses other motor systems to move the pickup across the disc surface in response to user commands. For example, the pickup must search rapidly across the disc as it reads data, or jump from one track to another. These functions are handled using control signals derived from the auto-tracking and auto-focus circuits; however, separate motors are used to move the pickup itself. Three beam pickups are mounted on a sled that moves across the disc surface. In many designs, linear motors move the pickup and position it to within capture range of the auto-tracking circuit, which takes control when the selected disc location is found. A spindle motor is used to rotate the disc with constant linear velocity. Thus the player must vary the disc speed depending on where the pickup is located on the disc surface—faster on inner diameters, and slower on outer diameters. This is accomplished with yet another servo loop; information from the data stream recovered by the laser pickup is used to determine correct rotating speed, and the spindle motor is regulated accordingly.

30.3.2 Data Decoding

The photodiode array and its processing circuits produce a signal resembling a series of high-frequency sinusoids called the *EFM signa*l. A collection of EFM waveforms (called an *eye pattern*) is shown in Fig. 30-9. The digital data can be recovered from the EFM signal if it can be determined when the signal crosses the zero axis, relative to the timing constraints created by the EFM encoding rules.

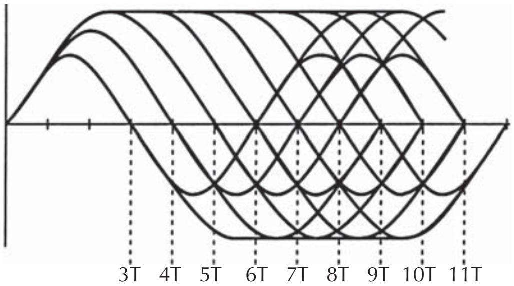


Figure 30-9. EFM eye pattern.

CD data decoding follows a procedure that essentially duplicates, in reverse order, the encoding process. The first data to be extracted from the signal is synchronization words. This information is used to synchronize the thirty three symbols of channel information in each frame, and a synchronization pulse is generated to aid in locating the zero crossing of the EFM pattern and to generate a transition at those points to produce a binary signal.

The EFM signal is demodulated so that every 17-bit EFM word is reconverted to 8-bits. Demodulation can be accomplished by logic circuitry or a look-up table. A buffer is used to remove the effect of disc rotational irregularities; data input to the buffer may be irregular in time but clocking ensures that the buffer output is precise. To guarantee that the buffer neither overflows nor underflows, a correction signal is generated and used to control the disc rotating speed.

Following demodulation, data is sent to a CIRC decoder for deinterleaving, error detection, and correction. The CIRC decoding process reverses the processing steps accomplished during encoding. The CIRC decoder accepts one frame of thirty-two 8-bit symbols; twenty four are audio symbols, and eight are parity symbols. One frame of twenty-four 8-bit symbols is output. The decoder utilizes parity from two Reed-Solomon decoders and deinterleaving. The first error correction decoder is designed to correct random errors, and to detect burst errors. It flags all burst errors, to alert the second error correction decoder.

Error concealment algorithms, employing interpolation and muting circuits, follow the CIRC decoder. Uncorrected words are detected through flags and dealt with, while valid data passes through unprocessed. Using error flags, the player's signal-processing circuits determine whether to output the data directly, to interpolate it, or to mute the sound.

For continuous errors, muting is employed as a last resort; invalid data passed on to the D/A converter could result in an audible click. Muting is accomplished by beginning attenuation many samples before the invalid data, smoothly muting the invalid data, and then smoothly restoring the signal level. This method of muting is often largely inaudible.

30.3.3 Signal Reconstruction

At the output stage, the digital data is converted to a stereo analog audio signal. This reconstruction requires low-pass filtering to suppress high-frequency image components and D/A conversion. An oversampling digital filter uses samples from the disc as input and then computes interpolation samples, digitally implementing the response of a low-pass filter. A transversal filter can be used to oversample (perhaps at an eight-times rate); image components appear at multiples of the new sampling rate. Because the separation between the baseband and sidebands is greater, a low-order analog filter can be used to remove the images. The type of oversampling filter found in CD players is an example of a wider class of FIR (finite impulse response) digital filters used in many applications. These kinds of filters use addition, multiplication, and delay elements to perform their tasks, and fall under a wider category of technology known as DSP (digital signal processing). The transversal filter used in CD players resamples and filters through interpolation. Resampling acts to increase the sampling rate; for example, in an eight-times oversampling filter, seven zero values are inserted for every data value output from the disc. This increases the sampling rate from 44.1 kHz to 352.8 kHz.

Interpolation is used to generate the values of intermediate sample points—for example, seven intermediate samples for each original sample. These samples are computed using coefficients derived from a low-pass filter response. In this way, when these samples are summed with other such samples, the output data stream corresponds to the sin $(x)/(x)$ impulse response processing of an ideal low-pass filter. Following this processing, the data is converted into a format appropriate for the type of D/A converter used in the player. In most CD players, sigma-delta D/A converters are used, employing techniques such as short word lengths, very high oversampling rates and noise shaping.

Also present in the audio output stage of every player is an audio deemphasis circuit. Some CDs are encoded with audio preemphasis characteristic. On playback, this is detected and deemphasis is automatically carried out, resulting in an improvement in signal-to-noise ratio.

30.4 Other CD Formats

The CD's small size, economy, robustness, and capacity make it an excellent music carrier. However, its utility is not limited to music playback. Other formats, including computer-based storage and recordable formats, have been derived from the original Red Book standard. In particular, the CD-ROM, CD-R, and CD-RW formats are widely used in computer applications as well as stand alone audio applications.

30.4.1 CD-ROM

The CD read-only memory (CD-ROM) standard, sometimes called the Yellow Book standard, is codified as the ISO/IEC 10149 standard. It is derived from the CD audio standard but defines a format for general data storage and is not tied to any specific application. Ninety-eight CD frames are summed to form a data block of 2352 bytes (24 bytes × 98 frames). Each disc holds 330,000 blocks. The first 12 bytes of a block form a synchronization pattern, and the next 4 bytes form a header field for time and address flags. The header contains three address bytes, represented as disc times, storing minutes, seconds, and block numbers within the second. The header also contains a mode byte; depending on the mode selected, the remaining 2336 bytes can store user data, or 2048 bytes of user data with extended error correction.

The mode byte identifies three modes and is used for two different data types, shown in Fig. 30-10. Mode 1 permits 2048 bytes of user data in each block. Each block contains 2 Kbytes (2 × 1024) of user data; 280 bytes are given to extended error detection and correction (EDC/ECC). A Mode 1 CD-ROM holds 682 million bytes of user information (333,000 blocks × 2048 bytes). Mode 2 gives the full 2336 bytes to user data. A CD-ROM bit stream is applied to conventional CD encoding so that CIRC, EFM, and other processing is applied. Mode 1 thus has two independent layers of error correction (EDC/ECC and CIRC) whereas Mode 2 uses only CIRC error correction.

Because of its extended error correction, EDC/ECC data independently supplements the CIRC error correction code applied to the frame structure, improving the error rate over that of audio CD. Mode 1 is employed for numerical data storage, which is more critical than audio data. In EDC/ECC encoding, a GF(2^8) Reed-Solomon product code (RS-PC) codes each block. It produces P and Q parity bytes with (26,24) and (45,43) code words respectively.

The CD-ROM/XA format is an extension to the Mode 2 standard and defines an XA data track that can contain diverse data such as computer, and compressed audio and video. However, CD-ROM/XA differs from CD-ROM Mode 2; XA provides a subheader that

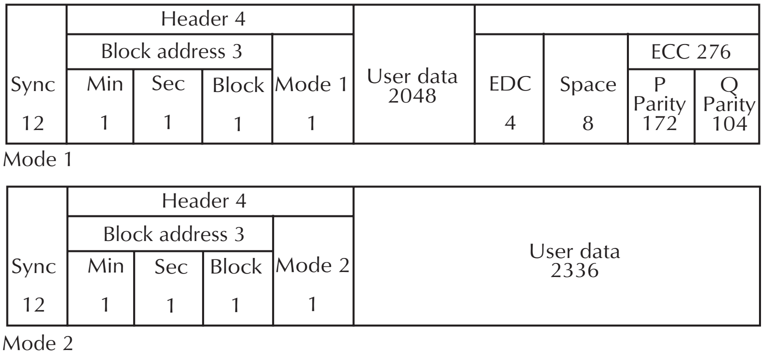


Figure 30-10. CD-ROM Mode 1 contains 2048 bytes of user data with extended error correction, and Mode 2 contains 2336 bytes of user data.

defines two types of blocks: Form 1 for computer data and Form 2 for compressed audio/video data. The former provides a 2048-byte user area, and the latter provides 2324 bytes. An XA track can interleave Form 1 and 2 blocks, but Red Book data cannot be placed in an XA track. Some products are dedicated to specific types of CD-ROM/XA discs; the Video CD is an example of this.

Hybrid audio/data CD formats such as CD Extra and Mixed Mode CD combine different format types (such as CD audio and CD-ROM/XA) on one disc. A CD Extra disc contains CD audio data in the first session, and CD-ROM-XA mode 2 data in the second session. In Mixed Mode CDs, ROM data is placed in track 1, and CD audio data is placed in subsequent tracks. To make sure an audio player does not access the ROM track, a *pregap* may be used so that ROM data is placed after the disc table of contents (TOC), but before the first music track. CD-ROM data is placed between Index 0 and Index 1 of Track 1, while the music starts at Track 1, Index 1. An audio player thus skips the data, starting playback at the first music track. However, the pregap area is not accessible to all drive software.

Unlike the CD audio standard, the CD-ROM standard does not stipulate how content is defined. Subsequently the ISO/DIS 9660 standard was devised; it specifies how computer data is placed on a CD-ROM; to read the data, the computer operating system must read the ISO 9660 file structure. Content on CD-ROM discs can be authored for multiple platforms; however, executable files can only run on the appropriate platform. For example, hybrid CD-ROM titles can be played on IBM and Macintosh platforms. The different data types are physically partitioned on the disc surface.

30.4.2 CD-R

The CD recordable (CD-R) format allows users to permanently record audio or other data to a CD. The format is technically named CD-WO (Write Once), as codified in the Orange Book Part II. CD-R discs that carry audio and nonaudio data prior to CD replication can be written with the PMCD (premastered CD) format; the disc contains index and other information. CD-R discs with up to 80 minutes of playing time (about 700 Mbytes) are available.

CD-R discs are physically different from Red Book CDs. CD-R discs are manufactured with a pregrooved spiral track with 0.6 µm width and 1.6 µm pitch; it guides the recording laser along the track. The pregroove is physically modulated with a ±0.03 µm sinusoidal wobble with a frequency of 22.05 kHz. Recorders use the wobble to control the disc CLV rotation speed. The 22.05 kHz groove wobble is also frequency modulated with a ±1 kHz signal; this creates an ATIP (absolute time in pregroove) clocking signal.

CD-R discs are manufactured on a polycarbonate substrate, and contain a metal (e.g., gold or silver) reflective layer, an organic dye recording layer, and a top protective layer. The recording layer is placed between the substrate and reflective layer as shown in Fig. 30-11. Together with the reflective layer it provides a reflectivity of about 73%. A writing laser with wavelength of 775 nm to 795 nm passes through the polycarbonate substrate and heats the recording layer to approximately 250°C, causing it to melt and/or chemically decompose to form a depression or mark in the recording layer. Simultaneously, the reflective layer is deformed. These depressions or marks have a decreased reflectivity. During readout, the same laser, reduced in power, is reflected from the data surface and its changing intensity is monitored.

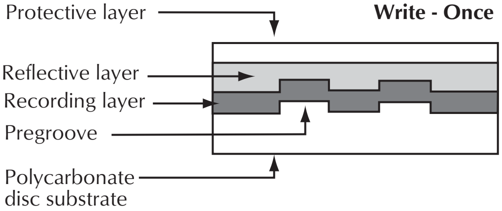


Figure 30-11. CD-R disc construction showing embedded recording layer.

Either cyanine or phthalocyanine organic dye polymers are often used for the recording layer. They are designed to absorb light at about 780 nm. Cyanine dye has a relatively broad range of sensitivity to light and is generally reliable in a wide range of recorders and laser powers and writing speeds. Phthalocyanine-based media are generally said to have greater longevity because it is less sensitive to ordinary light and stable. However, this lower sensitivity may result in a small power margin for the writing laser. Thus the writing speed and laser power must be more carefully controlled. In some cases, metallized azo dye is used as the recording layer in CD-R media. Organic dye layers are affected by aging. The dye layer will deteriorate over time because of oxidation, material impurities, or exposure to ultraviolet light. CD-R discs will play back on most CD audio players, but the reduced data layer reflectivity can cause playback incompatibility.

Two areas are written to the inner radius (22.35–23 mm) of CD-R discs, both inside the Red Book lead-in radius. The PMA (program memory area) contains data describing the recorded tracks, a temporary table of contents, and track skip information. When the disc is finalized, this data is transferred to the TOC. On the innermost radius, the PCA (power calibration area) is used by the recording laser to make an optimal power calibration test recording to determine proper laser recording power. A recording is complete when a lead-in area (with TOC), user data, and lead-out area are written. A maximum of 99 tracks can be recorded on a disc. Because the PMA and PCA areas are inside the normal lead-in radius, conventional CD players do not read them.

The CD-R standard defines both single-session and multisession recording (a session is a recording with lead-in, data, and lead-out areas). In single-session recording, sometimes called disc-at-once recording, an entire disc program is recorded without interruption. Track-at-once recording allows single or multiple tracks to be written in a session. Recorders using track-at-once can also write a single-session CD-R. In multisession recording, sessions can be recorded one or a few at a time. Tracks can be written singly and recording can be stopped after each track. Separate recording sessions are allowed, each with its own lead-in TOC, data, and lead-out areas. Track-at-once recorders allow both multisession and single-session recording. In track-at-once recording, multiple tracks can be written to a session, adding data one track at a time; no lead-in or lead-out is written until the session is closed. CD audio players can read only the first session on a multisession disc. A partially recorded disc can be played on the CD-R recorder but cannot be played on a CD audio player until the session ends when the final TOC and lead-out areas are recorded. Using the CD portion of the universal disk format (CD-UDF), CD-R recorders can perform packet writing; this allows small amounts of data to be written efficiently without high overhead. Data in a file can be appended and updated without rewriting the entire file.

30.4.3 CD-RW

The CD Rewritable (CD-RW) format allows data to be written and read, and erased and rewritten. The format is technically named CD-E and is described in the Orange Book Part III standard. A CD-RW drive can read, write, and erase CD-RW media, read and write CD-R media, and read CD-ROM and CD audio media. Thousands of rewrite cycles are possible. Any data can be written, including computer programs, text, pictures, video, audio, or other files. A CD-RW disc has five layers built on a polycarbonate substrate: a dielectric layer, a recording layer, another dielectric layer, a reflective aluminum layer, and a top acrylic protective layer, as shown in Fig. 30-12. As in CD-R, the writing and reading laser follows a pregroove spiral track. However, the CD-RW format employs a phase-change recording method, using materials that exhibit a reversible crystalline/amorphous phase change when recorded at one temperature and erased at another. In most cases, a high-reflectivity (crystalline) to low-reflectivity (amorphous) phase change is used to record data, and the reverse to erase. Data is recorded by heating an area of the crystalline layer to a temperature slightly above its melting point and cooled rapidly. The area is amorphous when it solidifies, and the decreased reflectivity is detected by a low power reading laser. Because the crystalline form is more stable, the material will tend to change back to this form. Thus when the area is heated to just below its melting temperature and cooled slowly, it returns to a crystalline state, erasing the data. In some cases, the recording layer comprises gallium antimonide and indium antimonide; other systems use tellurium alloyed with elements such as germanium and indium. The dielectric layers comprise silicon, oxygen, zinc, and sulfur; they control the optical response of the media and increase the efficiency of the laser by containing the heat in the recording layer. The dielectric layers also thermally insulate and protect the pregroove, substrate, and reflective layer.

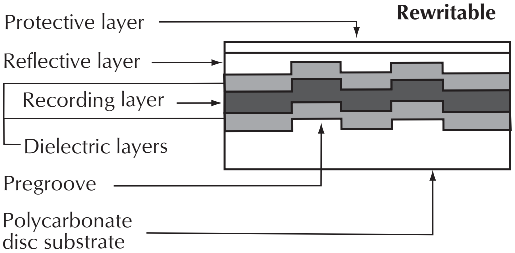


Figure 30-12. CD-RW disc construction showing embedded recording and dielectric layers.

The reflectivity of CD-RW discs is only about 15% (amorphous state) and 25% (crystalline). Discs will not play in most CD audio players or CD-ROM drives; however, many DVD players do play CD-RW discs. Multiread drives are capable of reading lower reflectivity CD-RW discs. They use an AGC (automatic gain control) circuit to boost the gain of the signal output from the photodiodes and compensate for the lower reflectivity and decreased signal modulation. CD-RW discs carry a code that identifies them as CD-RW discs to the player. CD-RW drives are commonly found as computer peripherals. Software supports track-at-once, disc-at-once, and multisession recording. When CD-RW discs are appropriately formatted, the CD-Universal Device Format (CD-UDF) specification permits easy file-by-file rewriting; for example, users can write to CD-RW discs with dragging and dropping.

30.5 Super Audio CD

The super audio CD (SACD) standard provides high-density storage to support two-channel CD and two-channel and multichannel SACD audio recordings. SACD recordings use 1-bit direct stream digital (DSD) coding with a high sampling frequency to achieve a frequency response to 100 kHz and a dynamic range of 120 dB in the 0 to 20 kHz band. Hybrid SACD discs can hold both a high-density DSD data layer (containing both a 5.1-channel mix and a stereo mix) as well as a Red Book compatible (44.1 kHz/16-bit) data layer. SACD players play back both SACD and CD discs. To achieve this, dual laser pickups operate at both the SACD 650 nm wavelength and the CD 780 nm wavelength. The SACD format also specifies a lossless coding algorithm known as direct stream transfer (DST); it uses an adaptive prediction filter and arithmetic coding to effectively double disc capacity. The SACD standard is described in the Scarlet Book.

30.5.1 SACD Specifications

SACD discs have a 12 cm diameter and 1.2 mm thickness, the same as a CD. Other specifications allow greater density; the laser wavelength is 650 nm, the lens numerical aperture (NA) is 0.60, the minimum pit/land length is 0.40 µm, and the track pitch is 0.74 µm. A single-layer SACD disc holds 4.7 Gbytes of data; this provides about 110 minutes of playing time for a two-channel stereo DSD recording. Several disc types are specified in the SACD format including single-layer, dual-layer, and hybrid disc constructions. The single-layer disc contains one layer of DSD content (4.7 Gbytes); the dual layer contains one or two layers of DSD content (8.5 Gbytes with two layers); and the hybrid disc is a dual layer disc that contains one inner layer of DSD content (4.7 Gbytes) and one outer layer of Red Book CD content (780 Mbytes) that can be played in ordinary CD players. In dual-layer discs, two 0.6 mm substrates are bonded together. There is only one data side in all implementations. A semireflective layer (20–40% reflective) covers the embedded inner data layer, and a fully reflective top metal layer (at least 70% reflective) covers the outer data surface. The outer data surface is protected by an acrylic layer and a printed label. Fig. 30-13 shows a hybrid disc and a dual pickup (650 and 780 nm) reading SACD and CD layers.

SACD players can play both SACD and CD discs (and hybrid SACD discs). CD data is passed to the digital filter and SACD data is applied to the DSD decoder. DSD data is output as a 1-bit signal and applied to a pulse density modulation processor. The data signal is converted to a complementary signal; each logical 1 creates a wide pulse and each logical 0 creates a narrow pulse. A current pulse D/A converter converts the voltage pulse output into a current pulse. This signal is applied to an analog low-pass filter to create the analog audio waveform. SACD recordings with DSD coding are not compatible with the DVD-Audio standard and its PCM coding. Some players may include decoders to accommodate both disc formats.

30.5.2 Direct Stream Digital Coding

SACD recordings employ direct stream digital (DSD) coding which uses 1-bit pulse density representation and sigma-delta modulation to code audio signals. Many A/D converters use sigma-delta techniques to sample the input signal at a high sampling frequency. The signal is applied to a decimation filter and quantized for output as a PCM signal at a nominal sampling frequency of 44.1 kHz (for CD) and up to 192 kHz (for

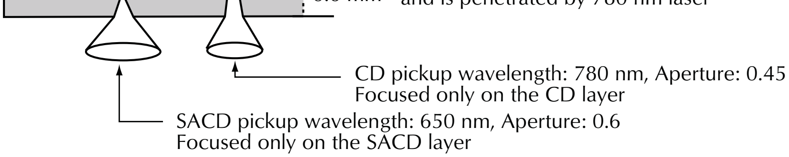

CD layer: Completely reflective
0.6 mm
SACD layer: Reflects 650 nm wavelength and is penetrated by 780 nm laser
0.6 mm
CD pickup wavelength: 780 nm, Aperture: 0.45
Focused only on the CD layer
SACD pickup wavelength: 650 nm, Aperture: 0.6
Focused only on the SACD layer

Figure 30-13. A hybrid SACD disc contains two data layers holding CD and SACD data.

DVD-Audio). Similarly, many D/A converters use oversampling to increase the sampling rate of the output signal and thus move the image spectra from the audio band. DSD coding uses a high sampling frequency, but does not require decimation filtering and multibit PCM quantization; instead, the original sampling frequency is retained. One-bit data is recorded directly on the disc. Moreover, DSD does not employ interpolation (oversampling) filtering during playback.

DSD uses sigma-delta modulation (SDM) and noise shaping. In a simple SDM encoder, the 1-bit output signal is used as a compensation signal. It is delayed by one sample and subtracted from the input analog signal using a negative feedback loop. If the input waveform rises above the value accumulated in the negative feedback loop during the previous sample, the converter outputs a logical 1. Similarly, if the waveform falls relative to the accumulated value, a logical 0 is output. The output pulses represent the magnitude of the input signal; pulse density modulation can be used. Because the integrator in the SDM encoder acts as low-pass filter, the low-frequency error content is reduced while the high-frequency content is increased. Higher-order noise shaping feedback filters can further decrease error in the audible range of frequencies. In principle, a low-pass filter can decode SDM signals, and also remove high-frequency noise resulting from noise shaping.

On SACD recordings, the DSD modulation employs a sampling frequency of 2.8224 MHz and each sample is quantized as a 2-bit word. Overall, the bit rate is thus four times higher than on a CD. In principle, the Nyquist frequency of DSD is 1.4112 MHz. However, to remove high-frequency noise introduced by noise shaping, some SACD players incorporate a 50 kHz low-pass filter (e.g., –3 dB at 50 kHz) for use with conventional power amplifiers and speakers. A 20 kHz low-pass filter is recommended when making SACD audio measurements. The 1-bit DSD signal can be converted to standard multibit PCM sampling rates.

30.6 DVD Disc Format

In its early development, DVD was envisioned as a consumer video disc playback system. Subsequent development expanded the scope of the standard. The resulting family of DVD optical disc formats encompasses video, audio, and computer applications, with both playback-only and recordable technologies. Because the scope of these applications far exceeded digital video, the original name of digital video disc was changed to digital versatile disc, but that name was never accepted. Instead, the format is simply called DVD.

Whereas the CD was originally designed exclusively for audio storage, and subsequently adapted to other applications, the DVD family was designed as a universal storage platform. The CD was designed to work with or without microcontrollers in the player. In contrast, DVD employs sophisticated microcontroller functions to read its file structure and interact with the disc and its contents. The CD was designed to play back a continuous stream of data. Thus, addressing was not provided; addressing capability was only subsequently developed for CD-ROM. In DVD, all data is addressable and randomly accessible; all DVD contents are essentially viewed as software data. Although its outer physical dimensions are identical, one DVD data layer provides about seven times the storage capacity of a CD. This increase is due to the shorter wavelength laser, higher numerical aperture, smaller track pitch, and other aspects.

The DVD family contains six DVD books: Book A is DVD-ROM (read only), Book B is DVD-Video, Book C is DVD-Audio, Book D is DVD-R (write-once), Book E is DVD-RAM (random access memory), and Book F is DVD-RW (rewritable). In each book, Part 1 defines the physical specifications, Part 2 defines the file system specifications, and subsequent parts define specific applications and extensions. For example, Part 3 defines the video application, Part 4 defines the audio application, and Part 5 defines the VAN extension. DVD-ROM video and audio discs use the same disc specifications and physical format as well as file system. DVD-R, DVD-RAM, and DVD-RW discs are more unique. The DVD format employs other specifications. For example, the DVD file system uses elements of the UDF, ISO 9660 and ISO 13346 specifications, and DVD-Video uses MPEG video coding and Dolby Digital (AC-3) audio coding.

The physical specifications for the DVD-ROM, DVD-Video, and DVD-Audio discs are identical and these read-only formats share disc construction, modulation code, error correction, etc. Discs are 120 mm or 80 mm in diameter and 1.2 mm in thickness, and have two bonded substrates with single or dual data layer per substrate. DVD discs use a pit/land structure to store data. The DVD track pitch is 0.74 µm. The track constant linear velocity (CLV) is 3.49 m/s on a single layer and 3.84 m/s on a dual layer. Minimum/maximum pit length is 0.40/1.87 µm (single layer) and 0.44/2.05 µm (dual layer). The laser beam used to read DVDs uses a wavelength of 635 nm or 650 nm. The objective lens has a numerical aperture of 0.6. A DVD layer can store 4.37 Gbytes (measured in 8-bit bytes) of data and multiple data layers provide greater capacity.

30.6.1 DVD Disc Manufacturing

A DVD thickness of 1.2 mm comprises two 0.6 mm substrates, bonded together with the data layers placed near the internal interface for greater protection. Thinner substrates are optically more resistant to tracking errors that result when a disc is slightly tilted relative to the laser pickup. The dual substrate construction allows manufacturing variants, yielding five types of playback-only discs: DVD-5 (single side, single layer), DVD-9 (single side, dual layer), DVD-10 (dual side, single layer), DVD-14 (dual side, mixed layers with single layer on one side and dual layer on the other side), and DVD-18 (dual side, dual layer). As the nomenclature loosely suggests, five disc capacities are supported: 4.37, 7.95, 8.75, 12.33, and 15.91 Gbytes (expressed in 8-bit bytes). When the average data output bit rate is 4.8 Mbps, the approximate playing times are DVD-5 (133 min), DVD-9 (241 min), DVD-10 (266 min), DVD-14 (375 min), and DVD-18 (482 min).

DVD manufacturing is similar to CD manufacturing. Following authoring, disc content is typically imaged on a hard drive disk, transferred to another medium such as Digital Linear Tape (DLT), and delivered to the disc mastering facility. A DLT Type III tape cartridge can hold up to 10 Gbytes of uncompressed data; with a transfer rate of 1.25 Mbytes/s, a 135 min program can be transferred in about 1 hour. A separate DLT is used for each physical disc layer. Alternatively, other media such as DVD-R or Exabyte may be used.

A single-layer, single-sided DVD-5 disc uses one substrate with a data surface and one blank substrate. Two substrates with data surfaces can be bonded together to form a single-layer, dual-sided DVD-10 disc; the disc is turned over to access the opposite layer. The DVD standard allows data to be placed on two layers in a substrate to create a dual-layer disc that is read from one side comprising a DVD-9 disc. The layers are separated by a clear resin and a very thin semitransparent (semireflective from 25% to 40%) sputtered layer of gold or silicon. Both layers are read from one disc side by moving the objective lens and focusing the reading laser on either layer. The beam either reflects from the lower semireflective layer or passes through it and reflects from the top reflective layer. Because the SNR and reflectivity of the interior layer are slightly reduced, the layer uses a faster linear velocity (3.84 m/s versus 3.49 m/s). Thus the pit length is longer (e.g., the minimum pit length is 0.44 µm versus 0.4 µm). The interior layer thus has less capacity than the top data layer.

In the manufacture of dual-sided discs, two polycarbonate substrates are independently formed and then bonded together using a hot-melt adhesive or UV-curable bonding. Dual-layer discs can be formed from two 0.6 mm substrates; one layer is fully metallized and the other is semireflectively metallized. The two substrates are then bonded together with a layer of UV-cured optically clear photopolymer. This technique can be used to manufacture single-sided discs (such as some DVD-9 discs). Alternatively, a single-layer substrate can be coated with a semitransparent layer followed by a layer of liquid photopolymer that is molded by a second stamper and hardened by exposure to ultraviolet light. After the layer is hardened, a fully reflective metal layer is applied and the substrate is bonded to a second substrate. This technique is used for some DVD-9 and DVD-18 discs. Construction of a dual-layer/dual-side DVD-18 disc is shown in Fig. 30-14.

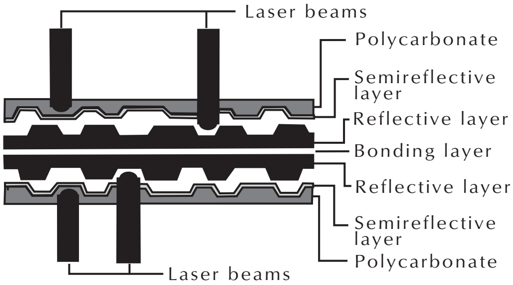


Figure 30-14. Construction of a dual-layer/dual-side DVD-18 disc.

30.6.2 DVD File Format and Coding

The DVD format is fundamentally computer-based with a file format defined for its applications. In particular, the DVD specification describes Universal Disc Format (UDF) Bridge, a file format specifically designed for optical disc storage. Read-only DVDs (DVD-ROM, DVD-Video, and DVD-Audio) use UDF for volume structure and file format, and UDF applies to the write-once and recordable disc formats. However, application-specific parameters are unique to both DVD-Video and DVD-Audio. UDF Bridge is a simplified version based on Part 4 of ISO/IEC 13346 and conforms to both UDF and ISO 9660 (the file format used in CD-ROM). UDF Bridge defines data structures such as volumes, files, blocks, sectors, CRCs, paths, records, allocation tables, partitions, and character sets, as well as methods for reading, writing, and other operations. It is a flexible, multiplatform, multiapplication, multilanguage, multiuser oriented format that has been adapted to DVD and is backward compatible to existing ISO-9660 operating system software. However, a DVD-Video or -Audio player supports only UDF and not ISO-9660.

In read-only DVD formats, data is stored in files within directories. DVD data is placed on a disc in physical sectors that run continuously without gap from the lead-in to the lead-out area. A DVD data sector comprises 2064 bytes, with 2048 bytes of main data and 16 header bytes; the latter comprises 4 bytes of identification (ID), 8 bytes of other data, and 4 bytes of error-detection code (EDC) data. The 4 bytes of identification data (ID) contain 1 byte of sector information and 3 bytes of sector number. A sync code is added to the head of every 91 bytes in the recording sector. This forms a physical sector. In all, 52 bytes of sync code is added. The 2048 bytes of user data is thus increased to 2418 bytes.

A Reed-Solomon Product Code (RS-PC) uses a combination of two Reed-Solomon codes (C1 and C2) as a product code. It differs from the CD's CIRC code. The two C1 and C2 product codes are (208,192) and (182,172) in length. Error correction is more challenging on a DVD because the pit size is smaller. In addition, because of the thin substrates, surface defects can more readily obscure the data surface. However, RS-PC is more powerful than the double error correction used in the CD-ROM format and provides improved error protection. RS-PC is also more efficient than CIRC in terms of overhead. In the DVD format, all disc types use the same level of error correction.

Read-only DVDs use EFMPlus modulation. It is an 8/16 RLL code and is similar to the EFM code used in CDs; for example, it uses the same minimum (2) and maximum (10) run length and represents logical 1 channel bits as pit/land or land/pit transitions and logical 0 channel bits as no transition. EFMPlus provides a 6% increase in user storage capacity compared to EFM because its coding is more efficient than EFM. Whereas EFM uses merging bits and a single lookup table and simple concatenation rules to suppress low-frequency content, EFMPlus does not require merging bits and uses a more sophisticated look-up method. The EFMPlus encoder defines four look-up tables each with 351 possible source words. In practice, the source codebook size is 344; seven possible words are discarded to allow for a unique 26-bit sync word. Of these, 256 words are used to code input data. The 88 surplus words are used as alternative channel representations to minimize the running digital sum value (DSV) and thus control low-frequency content.

In a DVD player, data passes through a buffer and is evaluated by a navigator/splitter that separates the bit stream into video, subpicture, audio, and navigational information. The video, subpicture, and audio data is decoded; for example, MPEG-2 video data is decoded as is Dolby Digital audio data. This can occur in a dedicated hardware chip or with software via a computer CPU. Navigational information is used by a controller for the user interface, while audio and video data is sent to the appropriate outputs.

30.7 DVD-Video

The DVD-Video format provides storage and playback of motion pictures or concert videos with multichannel soundtracks. The format was designed to provide the following: at least 133 minutes of digital video, approaching D1 broadcast picture quality, stereo or multichannel digital audio, multiple aspect ratios, up to 8 language soundtracks, up to 32 subtitles, parental control options, and copy protection.

In the DVD-Video format, data in a video disc is organized using the UDF Bridge file format. A DVD-Video zone and DVD-Other zone are defined under a root directory. In the DVD-Video zone, the VIDEO_TS directory (folder) contains menu and presentation data (video, audio, etc.). A Video Manager defines file types and organization of both video and audio data, and Video Title Set (VTS) subdirectories contain video and audio data files (such as MPEG-2 video and Dolby Digital audio). One Video Manager can contain up to 99 VTS subdirectories. Other computer data may be contained in the DVD-Other zone; this data may be used by DVD-ROM drives, and is ignored by DVD-Video players.

A DVD-Video VAN disc contains video-audio navigation data in a hybrid video-audio disc. VAN discs are video discs but they contain audio information that can be played on DVD-Audio players. Audio data is contained in an Audio Title Set, and video data in a Video Title Set. The Audio Manager and Video manager define file types and organize both audio and video data; both menu and program data is included. Using Link Info, a DVD-Audio player can play audio components of video contents.

The DVD-Video standard uses the MPEG-2 data compression algorithm to encode its video program. It employs the MPEG-2 Main Profile at Main Level protocol, also known as MP@ML. This is an intermediate level and below the high level sometimes used for DTV. However, MP@ML yields a high-quality picture that equals that of the professional CCIR-601 standard. The MPEG-2 video compression algorithm analyzes the video signal. Image data that is deemed redundant, not perceived, or marginally perceived is not coded or coarsely quantized. Analysis is carried out for both individual video frames (spatial reduction) and series of frames (temporal reduction). The video bit rate can be considerably reduced without significant degradation of the picture.

The video program is stored as 4:2:0 component video (Y, R-Y, B-Y) with progressive scan and picture resolution of 720 × 480 pixels. The picture quality of a particular DVD-Video title is primarily determined by the expertise of the picture encoding. The average output bit rate of a DVD-Video player is about 4.7 Mbps.

30.7.1 Audio Contents

Both stereo and multichannel soundtracks are accommodated in the audio portion of the DVD-Video standard. There can be 1 to 8 independent channels of linear PCM (LPCM), 1 to 6 channels of 5.1-channel Dolby Digital (AC-3), or 1 to 8 channels (5.1 or 7.1) of MPEG-2 AAC audio. A disc can also optionally employ DTS, SDDS, or other audio coding. Dolby Digital is the coding standard used for multichannel soundtracks in the United States (Region 1). The Dolby Digital sampling frequency is 48 kHz, the nominal output bit rate is 384 kbps, and the maximum bit rate is 448 kbps. Optionally, DTS codes multichannel audio data at a nominal bit rate of 1.4 Mbps. DTS can optionally be used to code 1 to 8 channels of audio, at sampling frequencies ranging from 8 kHz to 192 kHz. One DTS layer at a sampling frequency of 44.1 kHz can hold up to 74 min of 5.1-channel audio. MPEG-1 stereo audio is sampled at 48 kHz with a maximum bit rate of 384 kbps. MPEG-2 multichannel audio (up to eight channels) is also coded at 48 kHz; its maximum bit rate is 912 kbps. NTSC titles nominally use Dolby Digital, and PAL titles use MPEG-2 audio coding; however, PAL titles can optionally use Dolby Digital coding.

DVD-Video titles carry a redundant LPCM soundtrack employing sampling rates of either 48 or 96 kHz and word lengths of 16, 20, or 24 bits. These LPCM configurations are supported: 16/48 (up to eight channels), 20/48 (up to six channels), 24/48 (up to five channels), 16/96 (up to four channels), 20/96 (up to three channels), and 24/96 (up to two channels). The maximum LPCM bit rate is 6.144 Mbps on a DVD-Video. Various contents must be accommodated on a DVD-Video. For example, with an average video bit rate of 3.5 Mbps, there might be three audio soundtracks each at 0.384 Mbps, and 4 subtitles each at 0.01 Mbps, yielding a total bit rate of 4.692 Mbps. Thus, in this example, a DVD-5 would hold a 133 minute program.

Discs contain regional coding flags so players will only play certain regional discs. For example, a Region 2 (Europe and Japan) player will not play discs coded for the North American (Region 1) market. In this way, movie studios can control release of titles to different global markets. Regional coding of discs is optional; discs can carry multiple codes or no codes. Decoding circuitry is mandatory on all players.

DVD-Video optionally employ the Content Scrambling System (CSS) copy protection system. CSS-encoded content cannot be digitally copied because software keys needed to deencrypt the data are missing in any copy. Macrovision copy protection, similar to that used in set-top boxes, can be employed to prevent digital-to-analog copying of DVD-Video titles.

30.8 DVD-Audio

The DVD-Audio specification describes a high-fidelity audio storage medium supporting flexibility in the numbers of channels, sampling frequencies, word lengths, and other features such as video elements. DVD-Audio is principally used to code high-fidelity stereo and multichannel music programs using linear PCM (LPCM) data. Development of DVD-Audio was influenced by the International Steering Committee (ISC) representing the interests of the major record labels.

DVD-Audio was designed for compatibility with other DVD formats, some backward compatibility with the CD format, and to achieve improved sound quality and multichannel playback. Although the DVD-Video format can provide high-quality audio (such as six channels at 48 kHz/20-bit audio), its maximum audio bit rate of 6.144 Mbps cannot support the highest audio quality levels. Thus DVD-Audio's maximum bit rate was increased to 9.6 Mbps. However, six channels of 96 kHz/24-bit audio exceeds the maximum bit rate and high bit rates reduce playing time. Thus the Meridian Lossless Packing (MLP) lossless compression algorithm can be optionally employed to reduce bit rate, providing high fidelity and long playing time. This option allows storage of over 74 minutes of multichannel music on a single data layer. All DVD-Audio must contain an uncompressed or MLP-compressed LPCM version of the DVD-Audio portion of the program. For added compatibility with DVD-Video players, DVD-Audio may also include video programs with Dolby Digital, DTS, and/or LPCM tracks.

30.8.1 File Organization

Two types of DVD-Audio are defined. An Audio-Only disc contains primarily LPCM music content; it can optionally include still pictures (one per track), text information, and a visual menu. In an Audio-Only disc, data is contained in the DVD-Audio zone. The AUDIO_TS directory (folder) contains menu and presentation data. An Audio Manager defines file types and organizes audio and video data. Audio data, such as linear PCM, is contained in an Audio Title Set (ATS.)

An Audio with Video (AV) disc can contain motion video content formatted as a subset of the DVD-Video format. DVD-Audio players without video capability can play back the audio contents and audio components of video contents of DVD-Audio AV. They can play selected audio components on DVD-Video VAN. In an AV disc, audio data is contained in an Audio Title Set and video data in a Video Title Set. The Audio Manager and Video Manager define file types and organize audio and video data. Both menu and program data is included. The Audio Manager can control a subset of the DVD-Video data. Using Link Info a DVD-Audio player can play audio components of video contents. A DVD-Audio disc can be partially compatible with a DVD-Video player if the disc contains a stereo LPCM or Dolby Digital version of the album in the Video Title Set subdirectory of the disc. Universal DVD players can play all DVD-Audio and DVD-Video.

The presentation data for audio tracks is contained in AOB (Audio Object) files. Each AOB contains PCM data as well as optional audio data such as Dolby Digital. Optional nonaudio data such as still images are contained in ASV (Audio Still Video) files. The presentation data for video tracks is contained in VOB (Video Object) files. VOB files contain interleaved MPEG-2 data as well as audio data. Files needed to play back an audio track are located in the AUDIO_TS folder; files for video tracks are in the VIDEO_TS folder (also containing a VMG).

On many DVD-Audio, a track comprises one song. A disc can also contain up to nine groups per album (an album comprises one disc side). A group is essentially a playlist that contains up to 99 different tracks (each with up to 99 indices). A track may be included in more than one group. Users select a group and tracks within that group. This navigation is supported by the AMG (Audio Manager). The SAMG (Simple Audio Manager) is similar to a CD's TOC and contains a list of tracks (up to 314). Every disc includes a SAMG for track-based navigation. Simple players only recognize the SAMG and cannot recognize the AMG; these players have only two channel audio output and no video output. AMG players with video outputs read the AMG/AVTT (audio with video) section of the AMG. AMG players without video output read the AMG/AOTT (audio only) section. In this way, discs are compatible with players with widely different features. Fig. 30-15 summarizes the principal data elements found on a DVD-Audio.

Element	Outline of Contents
SAMG	Navigation information for simple audio player, which has only two-channel audio output
AMG	Information to navigate entire disc, may include optional text manager AMG menu video object set for visual menu
ASVS	Information to navigate still pictures Audio still video object set for still pictures
ATS	Information to navigate ATS Audio object set for audio data and optional RTI of audio tracks
VMG	Information to navigate video part
VTS	Information to navigate VTS Video object set for video/audio data of video tracks

Figure 30-15. Data elements found in a DVD-Audio.

30.8.2 Contents and Features

The DVD-Audio format supports a variety of coding methods and recording parameters. Optional audio coding methods include Dolby Digital, MPEG-1, MPEG-2 with/without extension bit stream, DTS, DSD, SDDS, and MLP. Linear PCM (LPCM) tracks are mandatory on all discs; all DVD-Audio players must support MLP decoding. Unlike some 5.1 channel systems (Dolby Digital, MPEG) the LPCM coding used in DVD-Audio does not band-limit the LFE channel; it is a full-bandwidth channel. DVD-Audio is a scalable format and gives flexibility to content providers. When LPCM coding is used, the number of channels (1 to 6), the word length (16, 20, 24 bit), and the sampling frequency (44.1, 48, 88.2, 96, 176.4, or 192 kHz) are all allowed. At the highest sampling frequencies of 176.4 kHz and 192 kHz, only two channel playback is possible. The audio coding options and the number of disc layers create a range of playback times. For example, a stereo LPCM program on a data layer might play for 258 minutes or 64 minutes, depending on its recording parameters. Similarly, different configurations of multichannel recordings will yield a range of playing times, as shown in Fig. 30-16. Use of MLP lossless compression, or lossy compression, increases playing times as well.

Audio channels are placed in two Channel Groups (CG). Examples of channel assignments are shown in Fig. 30-17. The grouping hierarchically lists mixes that

Audio Contents Combination	Channel Combination	Playback Time per Disc Side 8 cm Disc			
		Single Layer	Dual Layer	Single Layer	Dual Layer
2 channel only	48k/24-bit/2 ch	258 min	469 min	80 min	146 min
2 channel only	192k/24-bit/2 ch	64 min	117 min	20 min	36 min
2 channel only	192k/24-bi/2 ch	125 min	227 min	39 min	70 min
Multichannel Only	96k/24-bit/6 ch	86 min	156 min	27 min	48 min
2 ch and multichannel	96k/24-bit/2 ch +96k/24-bit/3 ch and 48k/24-bit/2 ch	76 min each	135 min each	23 min each	41 min each

Figure 30-16. Examples of playing times in DVD-Audio discs, not using MLP coding.

use the front L and R channels, front L, R and C channels, and the corner L, R, Ls, and Rs channels. The sampling frequency and word length of CG1 is greater than or equal to those of CG2. Generally, CG1 assignments are for front channels, and CG2 assignments are for rear channels. Channels can be assigned as groups of mono to six channels, and different word lengths and front and rear channels can use different sampling frequencies. For example, to reduce storage requirements, front channels could be coded at 24/96 and the rear channels coded at 16/48. The sampling frequencies must be related by a simple integer such as 48/96/192 kHz or 44.1/88.2/176.4 kHz.

Audio content can vary considerably. For example, a disc might use stereo LPCM audio for its selections. Another disc might contain one selection coded as multichannel LPCM and another coded as stereo LPCM. Another disc might contain one selection coded as stereo LPCM and another coded in an optional format such as Dolby Digital; advantageously, Dolby Digital tracks can be played in a DVD-Video player. Still another disc may include a DVD-Audio selection of up to six channels at 24/96 (possibly compressed with MLP), a stereo LPCM selection, and a Dolby Digital 5.1 channel selection on the DVD-Video portion.

DVD-Audio can employ the SMART (System Managed Audio Resource Technique) feature with LPCM tracks. Using SMART, a player can mix down a multichannel audio program to two channels for playback over a stereo system. The content provider controls the down-mixing by selecting one of sixteen coefficient tables. Each coefficient table defines level (0 to –60 dB), pan position, and phase; different tables can be used for each track in an Audio Title Set. With SMART, a separate stereo mix is not necessary on a multichannel disc, not wasting disc space. Use of SMART is optional on discs, but its support is mandatory in players.

		Channel Number					
		0	1	2	3	4	5
Mono/stereo playback	1	C					
	2	L	R				
Left front, Right front weighted	3	Lf	Rf	S			
	4	Lf	Rf	Ls	Rs		
	5	Lf	Rf	LFE			
	6	Lf	Rf	LFE	S		
	7	Lf	Rf	LFE	Ls	Rs	
	8	Lf	Rf	C			
	9	Lf	Rf	C	S		
	10	Lf	Rf	C	Ls	Rs	
	11	Lf	Rf	C	LFE		
	12	Lf	Rf	C	LFE	S	
	13	Lf	Rf	C	LFE	Ls	Rs
Front weighted	14	Lf	Rf	C	S		
	15	Lf	Rf	C	Ls	Rs	
	16	Lf	Rf	C	LFE		
	17	Lf	Rf	C	LFE	S	
	18	Lf	Rf	C	LFE	Ls	Rs
Corner weighted	19	Lf	Rf	Ls	Rs	LFE	
	20	Lf	Rf	Ls	Rs	C	
	21	Lf	Rf	Ls	Rs	C	LFE
				Channel Group 1		Channel Group 2	

Figure 30-17. Channel assignments using CG1 and CG2.

The DVD-Audio format uses optional content protection employing encryption and embedded watermark technology. The Content Protection for Pre-Recorded Media (CPPM) encryption code is stronger than that used in the DVD-Video format and has the capability to revoke, expire, or recover encryption keys. An optional CPPM watermark identifies content through unencrypted digital (and analog) links. It is not used in high-speed encrypted links and instead verifies copy status of unencrypted signals. The watermark is contained in the audio signal and is robust over analog and data-compressed transmission links.

30.8.3 Meridian Lossless Packing (MLP)

Meridian lossless packing (MLP) is an audio coding algorithm used to achieve lossless data compression. It reduces average and peak audio data rates and hence reduces storage capacity requirements. MLP packs audio data more efficiently, reducing file size without altering the contents. MLP offers other specific enhancements over PCM; whereas a PCM signal can be subtly altered by generation loss, transmission errors and other causes as it passes through a production chain, MLP can ensure that the output signal is exactly the same as the input signal by checking the MLP-coded file and confirming its bit accuracy. The compression achieved by MLP depends on the music being coded. Very approximately, it gives a 1.85:1 compression ratio; thus reducing the bit rate by almost 50%, doubling playing time with no loss of audio quality. For example, without compression, 96 kHz/24 bit audio requires 2.304 Mbps per channel. Thus a six channel recording would require 13.824 Mbps, exceeding DVD-Audio's 9.6 MHz maximum bit rate; thus LPCM cannot be used in the configuration. In contrast, MLP allows six-channel 96 kHz/24-bit recordings; it may achieve bandwidth reduction of 38% to 52%, reducing bandwidth to 6.6 to 8.6 Mbps, allowing a playing time of 73 to 80 minutes on a DVD-5 disc. In the two-channel stereo mode of 192 kHz/24-bit, MLP provides a playing time of about 117 minutes, versus a playing time of 74 minutes for LPCM coding.

Unlike lossy perceptual coding methods, MLP preserves bit-for-bit content of the audio signal. MLP provides less compression than lossy methods, the degree of compression depends on the audio signal content, and the output bit rate can continually vary according to signal conditions; however, a fixed data rate mode is provided. MLP is a mandatory coding option. Thus, all DVD-Audio players must support MLP decoding, but use of MLP on discs is optional for content providers. MLP may be used on a track-by-track basis. All of the DVD-Audio sampling frequencies are supported by MLP and quantization may be selected for 16 to 24 bits in 1-bit steps. MLP can code both stereo and multichannel signals simultaneously.

30.9 Other DVD Formats

The DVD-Video format is defined in Book B and DVD-Audio is defined in Book C. However, the DVD family also includes DVD-ROM (Book A), DVD-R (Book D), DVD-RAM (Book E), and DVD-RW (Book F). Books A, B, and C use the UDF Bridge file format while Books D, E, and F use the UDF format. The DVD-ROM, DVD-R, DVD-RAM, and DVD-RW formats are used primarily as computer peripherals or in professional authoring environments.

All DVDs are essentially DVD-ROMs, and all DVDs use the basic UDF format. Some DVD applications, such as DVD-Video, place specialized material in a specific place such as the DVD-Video zone. Content contained in the DVD-Other zone may be quite varied. DVD-ROM uses that provision for nonspecific storage, acting as a large capacity bit bucket formatted with UDF. DVD-ROM are playback-only media used to store data, software, games, etc. With appropriate software, DVD-ROM drives can play DVD-Video and DVD-Audio.

The DVD-R format offers write-once capability to permanently record data. DVD-Rs use a CLV wobbled pregroove to generate a carrier signal used for motor control, tracking and focus. DVD-Rs use pits and lands (known as land prepits) molded into land areas between grooves to encode the time address and other prerecorded signals. A cyanine organic dye recording layer may be used, with a 635 or 650 nm laser. The reading laser tracks the pregroove, but the light shines on the prepits peripherally to create a secondary signal that is extracted from the main signal. Discs can use the same reference velocity and track pitch as molded discs to achieve the same unformatted storage capacity. There are two parts to the DVD-R specification: DVD-R General and DVD-R Authoring; both yield discs playable on DVD-Video players.

DVD-R recorders perform an optimum power calibration (OPC) procedure to determine the correct laser writing power for particular discs, using a power calibration area (PCA) on discs to test laser writing power. A recording management area (RMA) saves calibration information, disc contents, and recording locations and remaining capacity information, recorder and disc identifiers for copy protection. The remainder of the disc comprises the information area containing the lead-in, data recordable area, and lead-out. The lead-in contains information on disc format, specification version, physical size and structure, minimum readout rate, recording density, and pointers to the location of the data recordable area where user data is recorded. The lead-out marks the end of the recording area. Both sequential (disc-at-once) and incremental writing can be performed. Once recorded, discs can potentially be played in DVD-ROM, DVD-Video, and DVD-Audio players. DVD+R is another write-once format using a dye recording layer and CLV rotation. Capacities of 4.7 and 8.5 (DL) Gbytes are available. DVD+Rs are generally compatible and can be played in many DVD-ROM, DVD-Video and DVD-Audio players.

The DVD-RW format allows data rewriting; the specification is an extension to the DVD-R format. Discs use a phase-change recording mechanism and a multilayer disc structure with dielectric layers above and below the recording layer. Data is recorded into a wobbled pregroove with CLV; relatively large data blocks are written. The recording layer may use a silver, indium, antimony, and tellurium compounded layer and allows perhaps 1000 writing cycles. Unlike dye-polymer technologies, phase-change recording is not wavelength-specific.

The DVD-RAM (random access memory) is a true random-access, nonsequential storage format. It uses a phase-change recording mechanism and a wobbled land and groove disc design. Data may be recorded on both planar surfaces of the groove and land; a wider track pitch is employed. This technique doubles disc capacity; deep grooves with steep walls are used to avoid crosstalk interference between adjacent data. Servos are used to switch the pickup's focus between the groove and land area on each revolution, and the tracking signal is inverted when the switch occurs. Discs contain preembossed pit areas (for every 2k sector) containing addressing header information and zoned constant linear velocity rotational control. DVD-RAM provides advanced error correction and defect management features. A disc allows perhaps 100,000 rewrite cycles and offers a high degree of stability for archiving integrity.

DVD+RW is a rewritable format that uses phase-change media, a wobbled pregroove, and CAV or CLV rotation, for either raw data transfer or faster data access. Data is recorded in the pregroove, not on the land. Data addresses are represented by modulation of the pregroove; this necessitates somewhat larger writing blocks. Over 100,000 rewrite cycles are possible. Both sequential and random access recording are supported.

30.10 HD DVD Format

New high-density disc formats are under continual development. HD DVD (high definition DVD) is one such format that can carry motion pictures in high-definition form, with picture quality greater than that of standard-definition DVD, and providing broadcast DTV high-definition quality. In addition, HD DVDs employ content protection that is more robust than that currently used in DVD.

The HD DVD format uses a 405 nm blue-violet laser and numerical aperture (NA) of 0.65. As with the DVD format, 12 cm diameter discs are formed from 0.6 mm substrates bonded together; data can be placed on one or both interface layers. Track pitch is 0.40 μm. A HD DVD-ROM holds 15 Gbytes on a single-layer disc, 30 Gbytes on a dual-layer disc, and 51 Gbytes on a triple-layer disc. A dual-side, single-layer disc holds 30 Gbytes and a dual-side, double-layer disc holds 60 GB. The structure of the HD disc is shown in Fig. 30-18. The single-speed transfer bit rate is 36 Mbps, and the double-speed bit rate is 72 Mbps. HD DVD movies have a maximum data bit rate of 36.55 Mbps (1×); maximum video bit rate is 28.0 Mbps. HD DVD supports the ISO 9660 and UDF optical disc file formats. HD DVD players can also play CDs and DVDs. HD DVD drives are available for the Xbox 360 and computer applications.

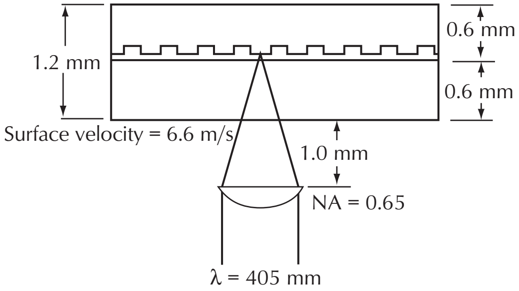


Figure 30-18. HD DVDs use two 0.6 mm substrates bonded together. A single-side disc is shown, but dual-side discs can be used.

The mandatory video codecs for HD DVD players are VC-1 (SMPTE 421M), MPEG-4 H.264 Advanced Video Codec (AVC), and MPEG-2. In practice, the majority of movie titles are coded with VC-1 at 1080p, with a minority coded with AVC. The HD DVD format supports a range of video resolutions, from low-resolution formats such as CIF and SDTV, to high-resolution formats such as HDTV at 720p, 1080i, and 1080p.

The mandatory audio codecs for players are Dolby Digital, Dolby Digital Plus, Dolby Digital EX, Dolby TrueHD (using MLP), DTS, and linear uncompressed PCM. Optional audio codecs include DTS-HD Master Audio and DTS-HD High Resolution Audio. Using these codecs, HD DVDs can contain up to eight channels of 24-bit/96 kHz audio, or two channels of 24-bit/192 kHz audio.

HD DVD optionally allows use of the Advanced Access Content System (AACS) for digital rights management, copy protection and content distribution control. AACS uses the Advanced Encryption Standard

(AES) to encrypt contents using one or more title keys. Content providers can revoke the decryption keys in individual players. Region coding is not used in the HD DVD format; any title can be played in any player. HD DVD uses the Microsoft HDi Interactive Format platform for interactive content on discs. HDi is based on existing protocols such as HTML, XML, CSS, SMIL, and JavaScript.

Alternative formats have been developed. The 3× DVD format uses a red laser; it yields approximately three times the storage capacity of DVD-Video; this format can hold high-definition content, but with shorter playing times. The HD REC format also stores high-definition content using a red laser and H.264/MPEG-4 AVC compression. The Combo hybrid disc is a dual-side disc (with up to two layers each) with a HD DVD layer and a DVD layer. A twin hybrid disc is a single-side disc with up to three layers, with either HD DVD or DVD content. Single-layer HD DVD-R and HD DVD-RW hold 15 Gbytes and dual-layer discs hold 30 GB. A single-layer HD DVD-RAM holds 20 GB. An experimental ten-layer disc would increase HD DVD storage capacity to 150 GB.

30.11 Blu-ray Disc Format

The Blu-ray disc system (also called BD) uses a 405 nm blue-violet laser and numerical aperture (NA) of 0.85 to achieve high storage capacity. Storage capacity is 25 Gbytes on a single-side, single-layer 12 cm disc. A single-side, dual-layer disc can hold 50 GB, or 9 hours of high-definition video. Track pitch is 0.32 µm, and the shortest pit length is 0.15 µm. The structure of the BD is shown in Fig. 30-19. The data layer is built on a 1.1 mm thick substrate and covered by a 0.1 mm spin-coated cover layer placed directly over the data layer and an optional top protection layer. The single-speed bit rate is 36 Mbps, the double-speed rate is 72 Mbps, the four times rate is 144 Mbps, and the six times rate is 216 Mbps. BD movies have a maximum data bit rate of 54 Mbps (1.5×); of this, the maximum video bit rate is 40 Mbps. A Blu-ray drive must operate at 1.5× speed to play BD movies. BD supports the ISO 9660 and UDF optical disc file formats. Although backward compatibility is possible, it is not required that BD players must also play CD and DVD-Video . BD drives are found in PS3 players and are available for computer applications.

The BD-ROM data layer is placed on the outer disc surface. The optical path is through a thin polymer layer that provides scratch resistance, not the substrate. Thus the substrate's optical characteristics are not critical—for example, birefringence is not a concern.

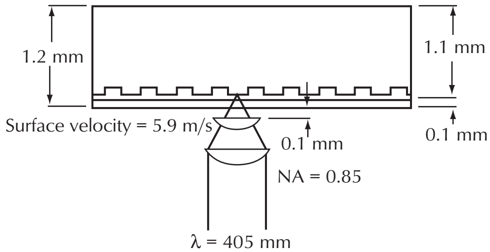


Figure 30-19. Blu-ray discs use a 1.1 mm substrate and a 0.1 mm protective layer. A single-layer disc is shown, but dual-layer discs can be used.

Because the objective lens is close to the data layer, optical aberration caused by disc tilt is limited. A 17PP modulation code is used, and a picket code with two Reed-Solomon codes is used for error correction.

The BD-ROM standard defines four player profiles. They describe functionality such as built-in persistent memory, local storage capability, secondary video decoder (for PiP), secondary audio decoder (for commentary and interactive content), virtual file system, and Internet connection capability. The four profiles are known as BD-Video (Grace Period Profile—Profile 1.0), Bonus View (Final Standard Profile—Profile 1.1), BD-Live (Profile 2.0), and BD Audio-Only (Profile 3.0).

The mandatory video codecs for BD-ROM players are VC-1 (SMPTE 421M), MPEG-4 H.264 Advanced Video Codec (AVC), and MPEG-2. The VC-1 and H.264 codecs are preferred; compared to MPEG-2, they provide greater compression and hence longer content run times, with similar quality. The BD format supports a wide range of video resolutions ranging from low to high resolution. The MPEG-2 transport stream is compatible with broadcast DTV.

Mandatory audio codecs for BD-ROM players are Dolby Digital, DTS, and linear uncompressed PCM (LPCM). A BD can hold up to up to eight channels of uncompressed LPCM audio. Optional audio codecs include Dolby Digital Plus, Dolby TrueHD (using MLP), DTS-HD Master Audio, and DTS-HD High Resolution Audio. Primary soundtracks must use a mandatory codec while secondary soundtracks can use either a mandatory or optional codec.

BD optionally allows use of the advanced access content system (AACS) for digital rights management, copy protection and content distribution control. AACS uses the advanced encryption standard (AES) to encrypt contents using one or more title keys. Title keys are formed from a media key, and the media's unique volume ID embedded on every disc. A broadcast encryption scheme is used such that each player has a unique set of decryption keys. Content providers can revoke the decryption keys in individual players. If a given key is compromised, that key can be revoked in future content, rendering it useless for decrypting future content. In addition, BD+ is a virtual machine resident in authorized players. It allows inclusion of executable security programs on BD. For example, programs can verify that AACS keys have not been altered, that hardware has not been tampered with, and can fix insecure systems. The BD-ROM Mark is cryptographic data that is physically stored in a manner that is different from other BD data; disc copies that do not contain the mark are not playable.

BDs can contain geographic region coding; content coded in a certain region will only play on that region's players. Region coding is optional; discs coded without a region code are playable in any player. There are three worldwide regions. Region A: North America, Central America, South America, Japan, Taiwan, North Korea, South Korea, Hong Kong, Southeast Asia; Region B: Europe, Greenland, French territories, Middle East, Africa, Australia, New Zealand. Region C: India, Bangladesh, Nepal, Mainland China, Pakistan, Russia, Central, South Asia. BDs use the Java software platform (a version called BD-J) for interactive content on discs. BD-J is a part of the Globally Executable MHP (GEM) standard; GEM is a version of the Multimedia Home Platform (DVB-MHP) standard.

A number of experimental BD architectures have been developed. They include a four-layer disc holding 100 GB, a six-layer disc holding 200 GB, and a ten-layer disc holding 250 GB. Alternative formats have been developed. The BD9 (Mini Blu-ray) format uses a red laser DVD to hold BD data. The disc is rotated at 3× speed to provide a minimum bit rate of 30.24 Mbps. Playing times are thus shorter than a conventional BD disc. The AVCREC format also uses red-laser DVD discs to hold BD content; H.264/MPEG-4 AVC compression is used. An experimental three-layer hybrid disc can hold both DVD-Video and BD data. Recordable (BD-R) and rewritable (BD-RE) Blu-ray disc formats are available, using phase-change technology. Dual-layer recordable discs are contemplated. Disc formats such as Blu-ray will further extend the opportunities of optical disc storage for professional and consumer applications.

References

1. K. C. Pohlmann, *The CD Handbook,* Second Edition, Madison, WI: A-R Editions Inc., 1992.

2. K. C. Pohlmann, *Principles of Digital Audio*, Fifth Edition, New York, NY: McGraw-Hill, 2005.

Bibliography

Blu-ray Disc, www.blu-ray.com.

G. Bouwhuis, et al., *Principles of Optical Disc Systems*, Adam Hilger Ltd., 1985.

M. G. Carasso, J. B. H. Peck, and J. P. Sinjou, "The CD Digital Audio System," *Philips Technical Review*, vol. 40, no. 6, 1982.

P. G. Craven and M. Gerzon, "Lossless Coding for Audio Discs," *Journal of the Audio Engineering Society*, vol. 44, no. 9, September 1996.

DVD Forum, www.dvdforum.org.

N. Fuchigami and T. Kuroiwa, "An Overview of the DVD-Audio Format," *Audio Engineering Society* preprint H-2, 9th Regional Convention of the Audio Engineering Society, Tokyo, June 23-25, 1999.

N. Fuchigami and T. Kuroiwa, B. H. Suzuki, "DVD-Audio Specifications," *Journal of the Audio Engineering Society*, vol. 48, no. 12, December 2000.

M. A. Gerzon, et al., "The MLP Lossless Compression System for PCM Audio," *Journal of the Audio Engineering Society*, vol. 52, no. 3, March 2004.

D. Goedhart and R. J. van de Plassche, and E. F. Stikvoort, "Digital-to-Analog Conversion in Playing a CD," *Philips Technical Review*, vol. 40, no. 6, 1982.

D. R. Guenette and D. J. Parker, “CD, CD-ROM, CD-R, CD-RW, DVD, DVD-R, DVD-RAM: The Family Album,” *EMedia Professional*, vol. 10, no. 4, April 1997.

J. P. J. Heemskerk and K.A.S. Immink, “CD System Aspects and Modulation,” *Philips Technical Review*, vol. 40, no. 6, 1982.

H. Hoeve, J. Timmermans and L. B. Vries, “Error Correction and Concealment in the CD System,” *Philips Technical Review*, vol. 40, no. 6, 1982.

IEC 908: *CD Digital Audio System*, 1987.

K. A. S. Immink, “The CD Story,” *Journal of the Audio Engineering Society*, vol. 46, no. 5, May 1998.

K. A. S. Immink, “The Digital Versatile Disc (DVD): System Requirements and Channel Coding,” *SMPTE Journal*, vol. 105, no. 8, August 1996.

Moving Pictures Expert Group, *DVD Resources*, www.mpeg.org/MPEG/dvd.html.

A. Nishio, et al., “Direct Stream Digital Audio System,” *Audio Engineering Society* preprint 4163, Copenhagen, May 1996.

D. J. Parker, “Writable DVD,” *EMedia Professional*, vol. 12, no. 1, January 1999.

K. C. Pohlmann, “The CD Formats: Technology and Applications,” *Journal of the Audio Engineering Society*, vol. 36, no. 4, April 1988.

J. R. Stuart, “Coding Methods for High-Resolution Recording Systems,” *Audio Engineering Society* preprint 4639, New York, September 1997.

B. H. Suzuki, “DVD-Audio Specifications,” *Proc. Audio Engineering Society 17*th *International Conference on High-Quality Audio Coding*, Florence, Italy, September 2-5, 1999.

J. Taylor, *DVD Demystified*, third edition, McGraw-Hill, 2006.

R. ten Kate, “Disc Technology for Super-Quality Audio Applications,” *Audio Engineering Society* preprint 4565, New York, September 1997.

J. Verbakel, et al., “Super Audio CD Format,” *Audio Engineering Society* preprint 4705, Amsterdam, May 1998.

L. B. Vries, et al, “The CD Digital Audio System: Modulation and Error Correction,” *Audio Engineering Society* preprint 1674, New York, October/November 1980.

This page intentionally left blank

Part 6

Design Applications

This page intentionally left blank

Chapter 31

DSP Technology

by Dr. Craig Richardson

31.1 Introduction . . . 1161
31.2 Digital Signal Processing . . . 1161
31.3 DSP Signals and Systems Theory . . . 1161
  31.3.1 Sequences . . . 1162
  31.3.2 Systems . . . 1162
  31.3.3 System Properties . . . 1164
    31.3.3.1 Linearity . . . 1164
    31.3.3.2 Time Invariance . . . 1164
    31.3.3.3 Causality . . . 1164
    31.3.3.4 Stability . . . 1164
  31.3.4 Linear Time-Invariant Systems . . . 1164
31.4 Frequency Domain Representation . . . 1165
31.5 The Z-Transform . . . 1166
31.6 Sampling of Continuous-Time Signals . . . 1167
  31.6.1 Continuous to Discrete Conversion . . . 1168
  31.6.2 Reconstructing the Continuous-Time Signal . . . 1169
  31.6.3 Sampling Theory . . . 1169
  31.6.4 Quantization . . . 1170
  31.6.5 Sample Rate Selection . . . 1171
31.7 Algorithm Development . . . 1172
31.8 Digital Signal Processors . . . 1172
  31.8.1 DSP Arithmetic . . . 1173
  31.8.2 Implementation Issues . . . 1174
  31.8.3 System Delay . . . 1174
  31.8.4 Choosing a DSP . . . 1176
31.9 Programming a DSP . . . 1176
31.10 Conclusion . . . 1177
References . . . 1178

This page intentionally left blank

31.1 Introduction

Over the past forty years, the field of digital signal processing (DSP) has grown from its origins as a collection of techniques for simulating the behavior of analog systems on digital computers into one of the most widely studied and universally used tools in modern technology. The use of DSP algorithms and implementations has become the rule rather than the exception, with applications in many areas such as music, communications, radar, sonar, image processing, robotics, seismology, meteorology, and applied physics. The remarkable growth of this discipline is largely due to two factors. First, DSP is a powerful problem-solving tool because it exploits the theoretical insights of discrete system theory to describe, analyze, and implement many interesting linear and nonlinear algorithms. Second, and more important, there is a special relationship between VLSI technology and DSP applications. The rapid development of digital integrated circuit technology has continually reduced the cost and increased the speed of the arithmetic operations necessary for DSP applications. In addition, DSP algorithms, which have demanding computational requirements but usually a very regular structure, are very well matched to the capabilities of VLSI. Integrated circuits are making complex DSP applications possible, and DSP applications have become a major motivating factor for building fast, complex integrated circuits. Perhaps the most visible embodiments of this phenomenon are the families of DSP microprocessors commonly called DSP chips. These chips have already had an immense impact on technology and are currently in the process of revolutionizing much of our industrial and technological base.

This chapter will introduce some of the important aspects of DSP technology including the fundamentals of DSP, the sampling process for converting analog signals to digital signals, the algorithm development process, and an introduction to programmable DSP devices. References are provided for finding additional information.

31.2 Digital Signal Processing

DSP is a technology and technique for analyzing and extracting information from signals, synthesizing signals, and manipulating signals. The acronym DSP is often used as both a noun and an adjective. DSP also often stands for digital signal processor—the actual microprocessor/computer that is used to implement the system. Common applications of DSP include cellular telephones, MP3 players, surround sound receivers, compact disc players, digital cameras, answering machines, and modems.

As with many disciplines, there are different perspectives and different layers of abstraction from which to explore DSP. For the purposes of this chapter, DSP will be approached and introduced from the theoretical, physical, and embedded software perspectives.

The theoretical perspective is concerned with the question “is something possible” and is built from fundamentals of DSP theory. This foundation includes linear system theory, complex number theory, and applied mathematics. The theoretical level provides a common language for DSP researchers to study and advance the state of the art.

The physical perspective is concerned with the devices that are used to implement DSP systems. These devices include the programmable digital signal processors that perform mathematical operations at a very high speed, and the details of converting an analog signal into a digital signal and then back to an analog signal.

The embedded software perspective is concerned with the actual software that makes the digital signal processors perform the desired tasks. This software is called *embedded* because it is executed internally on the DSP device and is only user accessible through some user interface, effectively hidden or embedded in the product, hiding the implementation details from the user.

31.3 DSP Signals and Systems Theory

The concepts of signals and systems are critical to an understanding of DSP. Signals can be a function of continuous time (i.e., analog) or of discrete time. Continuous-time signals have a signal value at any given instant of time while discrete-time signals only have a signal value at discrete instants of time. Values of discrete-time signals between the samples are determined by mathematically interpolating between the known sample values.

Signals represent the data that is to be processed. Examples include an audio file that needs to be compressed for low bit-rate storage or transmission or an image that will be searched for a particular object. A system is a transformation that maps an input signal (or multiple input signals) to an output signal (or multiple output signals)—i.e., the black box that maps inputs to outputs. In the music compression example, the output signal could be an MP3 file that was created by compressing an input signal. In the image example the output signal could simply be a yes/no decision along with positioning information. DSP systems are typically designed from simpler subsystems much like computer software is developed—subroutine by subroutine (one level of abstraction at a time). This section will introduce some fundamental systems and also introduce the useful properties that some systems possess.

31.3.1 Sequences

Discrete-time signals, also called sequences, are most often created by sampling analog, or continuous-time, signals. By sampling a continuous-time signal, a sequence of samples, really a sequence of numbers, can be processed and manipulated in a digital signal processor. Before going further into the sampling process, an introduction to signal and system theory will be presented, starting with discrete-time signals.

Discrete-time signals are represented mathematically as a sequence of numbers. The notation used will denote a sequence, x, as $x = \{x[n]\}$ where n is the index of the n^{th} element in the sequence. In terms of notation, $x[n]$ represents both the nth sample in the sequence and the entire sequence that is a function of n. The index, n, can range over all values from $-\infty$ to $+\infty$.

From a programming perspective, a sequence can be thought of as an infinitely large array of data indexed by an integer variable. In reality, an infinitely long array is not practical, so a sequence is usually represented as a continuous stream of data. Often it is assumed that the sequence starts at time = 0 ($n = 0$) and ends some finite time later ($n = M$).

There are several sequences that are fundamental building blocks of DSP systems. These are the unit impulse, the unit step sequence, and the sinusoid (cosine or sine). The unit impulse is a signal that has a value of 1 at index $n = 0$ and is 0 everywhere else as shown in Fig. 31-1. Mathematically this is denoted by

$$\delta[n] = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases} \tag{31-1}$$

Having defined the unit impulse, it is possible to represent a sequence $x[n]$ as a sum of delayed impulses that have a value of $x[k]$ at $n = k$. Mathematically this is formulated as

$$x[n] = \sum_k x[k]\delta[n-k] \tag{31-2}$$

which simply says that the value of $x[n]$ is the collection of its individual samples at time $n = k$.

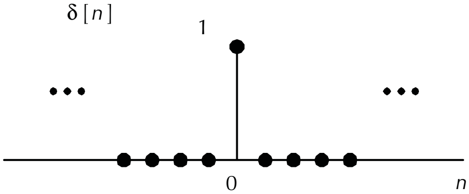


Figure 31-1. Unit impulse sequence has a value of 1 at $n = 0$ and is 0 everywhere else.

The unit step is a signal that starts at index 0 with value 1 and has value 1 for all positive indices as shown in Fig. 31-2. Mathematically, this is denoted by

$$u[n] = \begin{cases} 0, & n < 0 \\ 1, & n \geq 0 \end{cases} \tag{31-3}$$

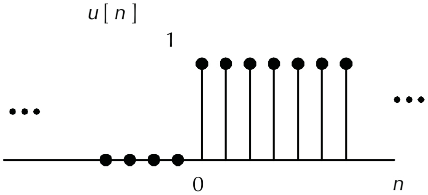


Figure 31-2. The unit step sequence has a value of 1 for $n \geq 0$ and is 0 everywhere else.

The cosine signal is a sinusoid of frequency ω and phase ϕ. An example of the cosine signal is shown in Fig. 31-3. Mathematically, the cosine signal is denoted by

$$\cos[n] = \cos(\omega n + \phi) \tag{31-4}$$

All sequences can also be represented by the numbers that are the sample values $x[n]$. Table 31-1 shows the sample values for the sequence in Fig. 31-4. Only the first sixteen sample values are listed because the sequence repeats itself after the 16th value ($x[15]$).

31.3.2 Systems

Systems transform input signals into output signals. Some commonly used systems include the ideal delay system that delays the output relative to the input and the moving average system that performs some simple low-pass filtering. Systems operate on a signal by operating on each sample individually or groups of samples at a time. For instance, multiplying a sequence by a constant can be implemented by multiplying each sample of the sequence by the constant. Similarly, the addition of two sequences is performed by adding the signals

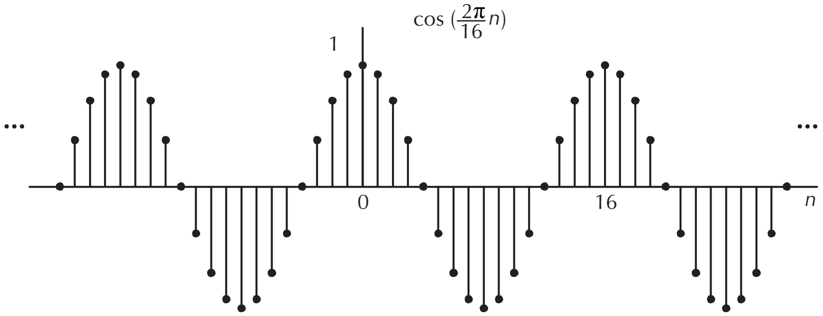


Figure 31-3. A cosine sequence of period 16. This particular cosine sequence is an infinite sequence of values that repeat with a period of 16 samples.

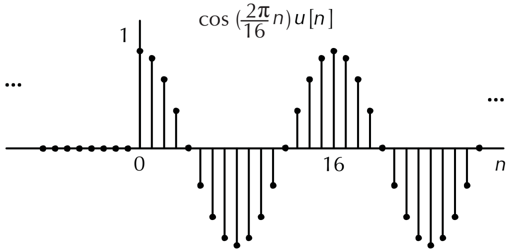


Figure 31-4. The product of the cosine sequence with the unit step sequence, $u[n]$. Notice that all the signal values for $n < 0$ are set to 0.

Table 31-1. The Values of the Signal $x[n]$ in Figure 31-4

$x[0]$	1.0000	$x[6]$	−0.7071	$x[12]$	0.0000
$x[1]$	0.9239	$x[7]$	−0.9239	$x[13]$	0.3827
$x[2]$	0.7071	$x[8]$	−1.0000	$x[14]$	0.7071
$x[3]$	0.3827	$x[9]$	−0.9239	$x[15]$	0.9239
$x[4]$	0.0000	$x[10]$	−0.7071	…	
$x[5]$	−0.3827	$x[11]$	−0.3827		

together on a sample-by-sample basis. Other systems, such as an MPEG audio compression system may operate on frames of data that have 1152 samples in each frame. The choice of whether to operate sample-by-sample or frame-by-frame is made by the system designer and algorithm developer.

A fundamental system is the ideal delay. The ideal delay system delays or advances a sequence by the delay amount. This system is defined by the equation

$$y[n] = x[n - n_d], -\infty < n < \infty \tag{31-5}$$

where,

n_d is an integer that is the delay of the signal.

The ideal delay system creates an output $y[n]$ by shifting the input signal, x, by n_d samples to the right when n_d is positive. This means that the value of the output signal $y[n]$ at a particular index n is the value of the input signal at index $n - n_d$. For example if the signal is delayed by three samples, then $n_d = 3$ and the output value $y[7]$ is equal to the value of $x[4]$—i.e., the value of $x[k]$ at $k = 4$ now appears at $y[j]$, $j = 7$. The system shifted the input signal three samples to the right as shown in Fig. 31-5.

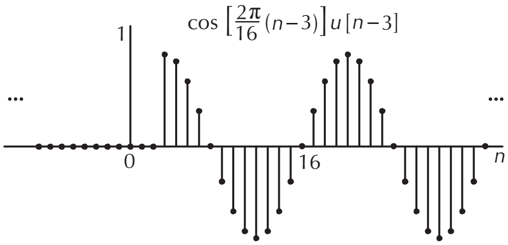


Figure 31-5. The cosine signal from Fig. 31-4 delayed by $n_d = 3$ samples. This delay shifts the sequence to the right by three samples.

The moving average system takes an average of the input signal over some window and then moves to the next sample and takes an average over the new window, etc. The general moving average system is defined by the equation below, where M_1 and M_2 are positive integers. It is called a moving average because to compute each output, $y[n]$, the filter must be moved to the next index and the average recomputed.

$$y[n] = \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} x[n-k] \tag{31-6}$$

The average sums values together starting M_1 samples forward from the current point and moving M_2 samples back from the current point and divides by the number of points that were summed together to form an average that smooths out the signal. The moving average is a digital filter that removes high frequency information through averaging.

31.3.3 System Properties

System properties are a convenient way to describe broad classes of systems. Important system properties include linearity, shift invariance, causality, and stability. These properties are important because they lead to a representation of systems that can be readily analyzed.

31.3.3.1 Linearity

A linear system is one where the output of a sum of linearly scaled input signals is equal to the sum of the linearly scaled output signals. Mathematically, a system, $T\{\bullet\}$, is linear when

$$y_1[n] = T\{x_1[n]\}$$

and

$$y_2[n] = T\{x_2[n]\}$$

Then

$$\begin{aligned} T\{ax_1[n] + bx_2[n]\} &= T\{ax_1[n]\} + T\{bx_2[n]\} \\ aT\{x_1[n]\} + bT\{x_2[n]\} &= ay_1[n] + by_2[n] \end{aligned} \tag{31-7}$$

This means that when the input to a linear system is a sum of signals, the output is the sum of the signals transformed individually.

As an example, consider a system that performs a scalar multiply $y[n] = \alpha x[n]$ (when $\alpha > 1$, $y[n]$ is a louder version of $x[n]$, and when $\alpha < 1$, $y[n]$ is a quieter version of $x[n]$). This system is linear because

$$\begin{aligned} y[n] &= \alpha(ax_1[n] + bx_2[n]) \\ &= (\alpha ax_1[n] + \alpha bx_2[n]) \end{aligned}$$

An example of a nonlinear system would be a compressor/limiter because the output of a compressor/limiter to a sum of signals is generally not equal to the sum of the compressor/limiters applied to the signals individually.

31.3.3.2 Time Invariance

A time-invariant system is one where a delay in the input signal causes the output to be delayed by the same amount. Mathematically, a system, $T\{\bullet\}$, is time invariant if when $y[n] = T\{x[n]\}$ then

$$T\{x[n-N]\} = y[n-N] \tag{31-8}$$

When the input, $x[n]$, to a linear system is delayed, the output, $y[n]$, is delayed correspondingly. There is no absolute time reference associated with the system. The combination of time invariance and linearity makes the design and analysis of a large class of DSP theory and applications much simpler due to the convolution operation and Fourier analysis tools.[1]

31.3.3.3 Causality

A causal system is one where the output of the system at a given time only depends on the present and past values of the input signal. No future data can be required to produce an output signal at the present time in a causal system. In the moving average system of Eq. 31-6, the system is causal only if $M_1 = 0$.

31.3.3.4 Stability

A system is bounded input/bounded output stable if and only if every bounded input sequence produces a bounded output sequence. A sequence is bounded if each value in the sequence is less than infinity. For real applications, stability is critically important because a system would stop operating properly should it ever become unstable.

31.3.4 Linear Time-Invariant Systems

When the linearity property is combined with the time-invariance property to form a linear time-invariant (LTI) system, then the analysis of systems is very straightforward. Because a sequence can be represented as a sum of weighted delayed impulses as shown in Eq. 31-2, and an LTI system response is the sum of the component responses of the sequence components as shown in Eq. 31-7, the response of an LTI system is completely determined from its response to an impulse. Since an input signal can be represented as a collection of delayed and scaled impulses, the response to the full sequence is known. The response of a system to an impulse is commonly referred to as the *impulse response* of the system. Mathematically,

$$x[n] = \sum_k x[k]\delta[n-k]$$

i.e., the sequence $x[n]$ is a sum of scaled and delayed impulses. If $h_k[n] = T\{\delta[n-k]\}$, i.e., the system response to the delayed impulse at $n = k$, then the output $y[n]$ can be formed as

$$\begin{aligned} y[n] &= T\{x[n]\} \\ &= T\left\{\sum_k x[k]\delta[n-k]\right\} \\ &= \sum_k x[k]h_k[n] \end{aligned} \qquad (31\text{-}9)$$

If the system is also time invariant, then $h_k[n] = h[n - k]$, and the output $y[n]$ is given by

$$\begin{aligned} y[n] &= \sum_k x[k]h[n-k] \\ &= \sum_k h[k]x[n-k] \end{aligned} \qquad (31\text{-}10)$$

This representation is known as the convolution sum and is commonly written as $y[n] = x[n] \bullet h[n]$. The convolution system takes two sequences, $x[n]$ and $h[n]$, and produces a third sequence $y[n]$. For each value of $y[n]$, the computation requires multiplying $x[k]$ by $h[n - k]$ and summing over all valid indices for k where the signals are non-zero. To compute the output $y[n + 1]$, move to the next point, n + 1, and perform the same computation. The convolution is an LTI system and is a building block for many larger systems.

As an example, consider the convolution of the sequences in Fig. 31-6 where $h[n]$ has only three non-zero sample values and $x[n]$ is a cosine sequence that has non-zero sample values for $n \geq 0$.

The computation of

$$y[n] = \sum_{k=0}^{2} h[k]x[n-k]$$

is performed as follows. Values of $x[n]$ for $n < 0$ are 0. Only the computation for the first three output samples are shown.

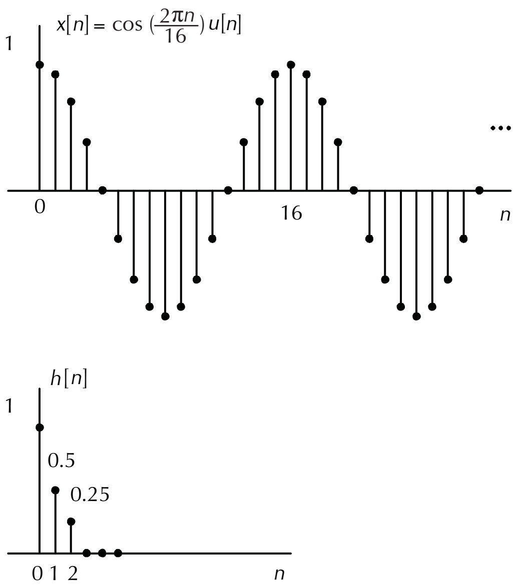


Figure 31-6. A convolution example with two sequences. $x[n]$ is the same signal from Fig. 31-4 with values shown in Table 31-1, and $h[n]$ has the values shown above.

$$\begin{aligned} y[0] &= h[0]x[0] + h[1]x[-1] + h[2]x[-2] \\ &= 1.0 \\ y[1] &= h[0]x[1] + h[1]x[0] + h[2]x[-1] \\ &= 1.4239 \\ y[2] &= h[0]x[2] + h[1]x[1] + h[2]x[0] \\ &= 1.4190 \end{aligned}$$

The result of the convolution is shown in Fig. 31-7 and has the sample values shown in Table 31-2.

31.4 Frequency Domain Representation

Having defined an LTI system, it is possible to look at the signal from the frequency domain perspective and understand how a system changes the signals in the frequency domain. The frequency domain represents signals as a combination of various frequencies from low frequency to high frequency. Each time-domain signal has a representation as a collection of frequency components where each frequency component can be thought of as sinusoids or tones. Sinusoids are important because a sinusoidal input to a linear time-invariant system generates an output of the same frequency but with amplitude and phase determined by the system. This property makes the representation of signals in terms of sinusoids very useful.

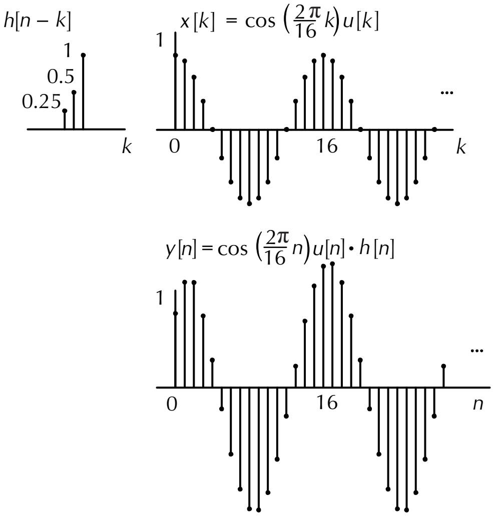


Figure 31-7. The output, y[n], from the convolution of *x*[*n*] and *h*[*n*] in Fig. 31-6.

Table 31-2. The Result of the Convolution in Fig. 31-7

$y[0]$	1.0000	$y[11]$	–0.9672	$y[22]$	–0.8984
$y[1]$	1.4239	$y[12]$	–0.3681	$y[23]$	–1.3731
$y[2]$	1.4190	y[13]	0.2870	$y[24]$	–1.6387
$y[3]$	0.9672	$y[14]$	0.8984	$y[25]$	–1.6548
$y[4]$	0.3681	$y[15]$	1.3731	$y[26]$	–1.4190
$y[5]$	–0.2870	$y[16]$	1.6387	$y[27]$	–0.9672
$y[6]$	–0.8984	$y[17]$	1.6548	$y[28]$	–0.3681
$y[7]$	–1.3731	$y[18]$	1.4190	$y[29]$	0.2870
$y[8]$	–1.6387	$y[19]$	0.9672	$y[30]$	0.8984
$y[9]$	–1.6548	$y[20]$	0.3681	…	
$y[10]$	–1.4190	$y[21]$	–0.2870		

As an example, assume an input signal $x[n]$ is defined as $x[n] = e^{j\omega n}$ —i.e., a complex exponential (Euler's relationship from complex number theory that states that $e^{j\omega n} = \cos(\omega n) + j\sin(\omega n)$, where ω is the radian frequency that ranges from $0 \le \omega \le 2\pi$), then using the convolution sum of

$$y[n] = \sum_k h[k]x[n-k]$$

generates

$$y[n] = \sum_k h[k]e^{j\omega(n-k)} \tag{31-11}$$

$$y[n] = e^{j\omega n}\left(\sum_k h[k]e^{-j\omega k}\right) \tag{31-12}$$

By defining

$$H(e^{j\omega}) = \sum_k h[k]e^{-j\omega k}$$

we have

$$y[n] = H(e^{j\omega})e^{j\omega n}$$

where,

$H(e^{j\omega})$ represents the phase and amplitude determined by the system.

This shows that a sinusoidal (or, in this case, the complex exponential) input to a linear time invariant system will generate an output that has the same frequency but with an amplitude and phase determined by the system.

$H(e^{j\omega})$ is known as the frequency response of the system and describes how the LTI system will modify the frequency components of an input signal. The transformation

$$H(e^{j\omega}) = \sum_k h[k]e^{-j\omega k}$$

is known as the Fourier transform of the impulse response, $h[n]$. If $H(e^{j\omega})$ is a low-pass filter, then it has a frequency response that attenuates high frequencies but not low frequencies—hence it passes low frequencies. If $H(e^{j\omega})$ is a high-pass filter, then it has a frequency response that attenuates low frequencies but not high frequencies.

In many instances it is more useful to process a signal or analyze a signal from the frequency domain than in the time domain either because the phenomenon of interest is frequency based or our perception of the phenomenon is frequency based.

An example of this is the family of MPEG audio compression standards that exploits the frequency properties of the human auditory system to dramatically reduce the number of bits required to represent the signal without significantly reducing the audio quality.

31.5 The Z-Transform

The Z-transform is a generalization of the Fourier transform that permits the analysis of a larger class of systems than the Fourier transform. In addition, the analysis of systems is easier due to the convenient notation of the Z-transform.[1] The Fourier transform is defined as

$$X(e^{j\omega}) = \sum_{k} x[k]e^{-j\omega k}$$

while the Z-transform is defined as

$$X(z) = \sum_{k} x[k]z^{-k}$$

When working with linear time invariant systems, an important relationship is that the Z-transform of the convolution of two sequences is equal to the multiplication of the Z-transforms of the two sequences, i.e., $y[n] = x[n]*h[n] \Leftrightarrow Y(z) = X(z)H(z)$. $H(z)$ is referred to as the system function (a generalization of the transfer function from Fourier analysis).

A common use of the Z domain representation is to analyze a class of systems that are defined as linear constant-coefficient difference equations that have the form of

$$\sum_{k=0}^{N} a_k y[n-k] = \sum_{k=0}^{M} b_k x[n-k] \qquad (31\text{-}13)$$

where,

the coefficients a_k and b_k are constant (hence the name constant coefficient).

This general difference equation forms the basis for both finite impulse response (FIR) linear filters, and infinite impulse response (IIR) linear filters. Both FIR and IIR filters are used to implement frequency selective filters (e.g., high-pass, low-pass, bandpass, bandstop, and parametric filters) and other more complicated systems.

FIR filters are a special case of Eq. 31-13, where except for the first coefficient, all the a_k are set to 0, leading to the equation

$$y[n] = \sum_{k=0}^{M} b_k x[n-k] \qquad (31\text{-}14)$$

The important fact to notice is that each output sample $y[n]$ in the FIR filter is formed by multiplying the sequence of coefficients (also known as *filter taps*) by the input sequence values. There is no feedback in an FIR filter—i.e., previous output values are not used to compute new output values. A block diagram of this is shown in Fig. 31-8 where the z^{-1} blocks are used to denote a signal delay of one sample (i.e., the Z-transform of the system $h[n] = \delta[n-1]$).

An IIR filter contains feedback in the computation of the output $y[n]$—i.e., previous output values are used to create current output values. Because of this feedback, IIR filters can be created that have a better frequency response (i.e., steeper slope for attenuating signals outside the band of interest) than FIR filters for a given amount of computation. However, most DSP architectures are optimized for computing FIR filters—i.e., multiplying and adding signals together continuously—so the choice of which filter style to use will depend on the particular application.

31.6 Sampling of Continuous-Time Signals

The most common way to generate a digital sequence is to start with a continuous-time (analog) signal and create a discrete-time signal. For example, speech signals are continuous-time signals because they are continuous waves of acoustic pressure. A microphone is the transducer that converts the acoustic signal into a continuous-time electric signal. In order to process this signal digitally, it is necessary to convert this signal into the digital domain. Finally, after processing, it is often necessary to convert the discrete-time signal back into a continuous-time signal for playback through a loudspeaker system.

The process of converting an analog signal to a digital signal is often be modeled as a two-step process, as shown in Fig. 31-9, of converting a continuous-time signal to a discrete-time signal (with infinite resolution of the amplitude) and then quantizing the discrete-time signal into finite precision values (creating the digital sequence) that can be processed by a computer.[1] The process of converting the continuous-time signal into a discrete-time signal will be introduced, and then quantization will be reviewed. The quantization step is necessary to create a sample value that has a data word size that is compatible with the arithmetic capabilities of the target DSP. All real-world analog-to-digital converters (A/Ds) perform both the sampling and quantization process internal to the A/D device, but it is useful to discuss the subsystems separately because they have different significance and design trade-offs.

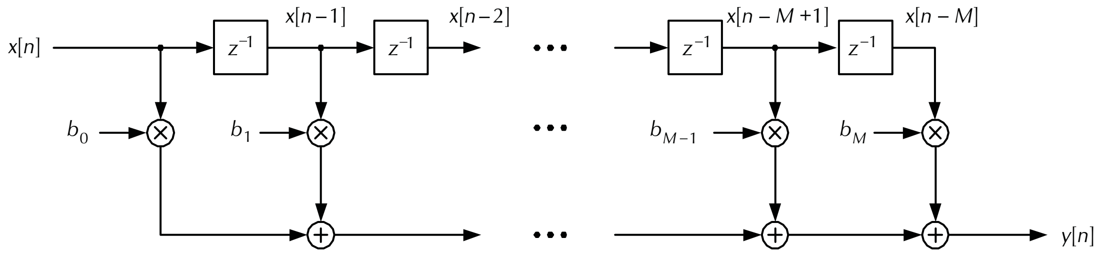


Figure 31-8. A block diagram of an FIR system where the input $x[n]$ is fed into a system that multiplies the delayed input signal with the filter coefficients b_k and sums the results together to form the output $y[n]$.

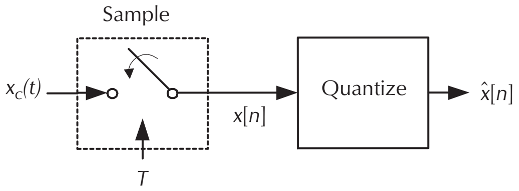


Figure 31-9. Analog-to-digital conversion can be thought of as a two-step process: converting a continuous-time signal to a discrete-time signal, $x[n]$, followed by quantizing the sample to create the digital sequence.

31.6.1 Continuous to Discrete Conversion

The most common method for converting a continuous-time signal, $x_c(t)$, into a discrete-time signal, $x[n]$, is to uniformly sample the signal every T seconds with the equation

$$x[n] = x_c(nT), -\infty < n < \infty \tag{31-15}$$

This generates a sequence of samples, $x[n]$, where the value of $x[n]$ is the same as the value of $x_c(t)$ whenever $t = nT$—i.e., at each sampling interval T. $1/T$ is known as the *sampling frequency* and is usually expressed in Hertz or cycles per second.

Mathematically, when a continuous-time signal is sampled, the resulting signal has a frequency response that is related to the underlying continuous-time signal frequency response and the sampling rate. As shown next, this has significant ramifications for how often the signal must be sampled in order for the digital sequence to be reconstructed into an analog signal that accurately represents the original signal.

The sampling process will be analyzed in the frequency domain where it will be assumed that a band limited signal, $x_c(t)$, is to be sampled periodically with sample period T. A band-limited signal is one that has no signal energy higher than a particular frequency, Ω_N, as shown in Fig. 31-10, where Ω represents the frequency axis of the signal. The reason the signal is assumed to be band limited is to prevent frequency aliasing, as will be evident shortly. The assumption of being band limited is significant although generally easily realizable in real-world systems.

The sampling of the continuous-time signal, $x_c(t)$, generates a signal, $x_s(t)$, from equation

$$x_s(t) = \sum_{n=-\infty}^{\infty} x_c(nT)\delta(t - nT) \tag{31-16}$$

$x_s(t)$ is the collection of values of $x_c(t)$ at the sampling interval of T. A convenient representation of this signal is as a collection of delayed and weighted impulse functions. The amplitude is the value at the sampling instant and the samples are spaced out by the sampling period T. The process can be analyzed in the frequency domain by first representing the Fourier transform of the impulse sequence as a sequence of impulses in the frequency domain[6]. This means that a sequence of equally spaced impulses in the time domain have a frequency representation that is a sequence of equally spaced impulses in the frequency domain, spaced by the sampling frequency $2\pi/T$. This is shown as

$$S(j\Omega) = \frac{2\pi}{T}\sum_{k=-\infty}^{\infty} \delta(\Omega - k\Omega_s) \tag{31-17}$$

where,

$\Omega_s = 2\pi/T$ is the sampling frequency in radians/second.

The Fourier transform of the sampled signal, $x_s(t)$, becomes

$$X_s(j\Omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_c(j(\Omega - k\Omega_s)) \tag{31-18}$$

Now the frequency response of the sampled continuous-time signal becomes a collection of shifted copies of the original frequency response of the analog signal $X_c(j\Omega)$. Fig. 31-10 shows the frequency response of $X_c(j\Omega)$, the impulse train, $S(j\Omega)$, and the resulting frequency response of the sampled signal, $X_s(j\Omega)$.

This frequency response, $X_S(j\Omega)$, can also be interpreted as the convolution in the frequency domain between the frequency response of the continuous-time signal and the frequency response of the impulse train, $S(j\Omega)$.

$$X_s(j\Omega) = \frac{1}{2\pi}X_c(j\Omega)*S(j\Omega) \tag{31-19}$$

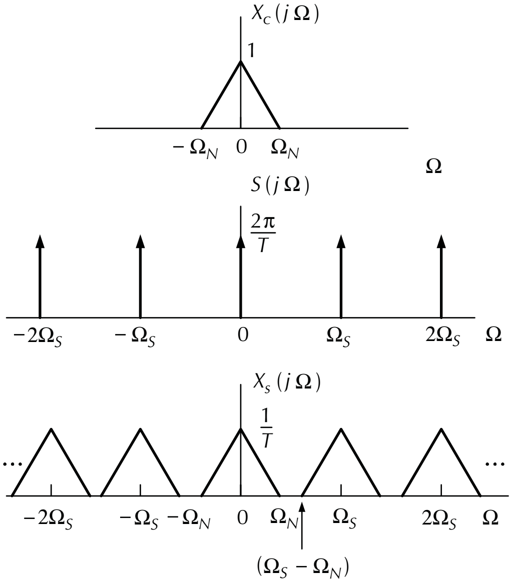


Figure 31-10. The frequency response of the analog signal, $X_c(j\Omega)$, the sampling function, $S(j\Omega)$, and the resulting frequency response of the sampled signal, $X_s(j\Omega)$.

From Fig. 31-10 it can be seen that as long as the sampling frequency minus the highest frequency is greater than the highest frequency, $\Omega_S - \Omega_N > \Omega_N$, the frequency copies do not overlap. This condition can be rewritten as $\Omega_S > 2\Omega_N$, which means that the sampling frequency must be at least twice as high as the highest frequency in the signal. If the sampling frequency is less than the highest frequency in the signal, $\Omega_S < 2\Omega_N$, then the frequency copies overlap as shown in Fig. 31-11. This overlap causes the frequencies of the adjacent spectral copies to be added together, which results in the loss of spectral information. It is impossible to remove the effects of aliasing once aliasing has happened. The overlap is caused because the sampling frequency, Ω_S, is not high enough relative to the highest frequency in the continuous-time signal $X_c(j\Omega)$. As shown above, the sampling frequency must be at least twice as high as the highest frequency in the continuous-time signal in order to prevent this overlap, or aliasing, of frequencies.

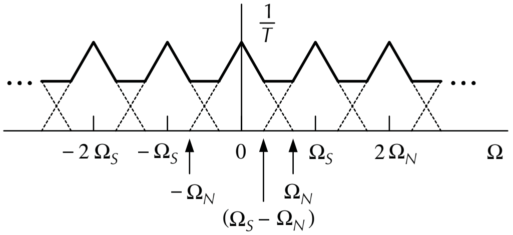


Figure 31-11. Sampling where the sampling frequency, Ω_S, is less than twice the highest frequency, Ω_N.

31.6.2 Reconstructing the Continuous-Time Signal

As seen from sampling a continuous-time signal, if the signal is not sampled fast enough, then the resulting frequency response of the sampled signal will have overlapping copies of the frequency response of the original signal. Assuming the signal is sampled fast enough (at least twice the bandwidth of the signal), the continuous-time signal can be reproduced by simply removing all of the spectral copies except for the desired one. This frequency separation can be performed with an ideal low-pass filter with gain, T, and cut-off frequency, Ω_C, where the cut-off frequency is higher than the highest frequency in the signal as well as the frequency where the first frequency replica starts,—i.e., $\Omega_N < \Omega_C < \Omega_S - \Omega_N$. Fig. 31-12 shows the repeated frequency spectrum and the ideal low-pass filter. Fig. 31-13 shows the result of applying the low-pass filter to $X_S(j\Omega)$.

31.6.3 Sampling Theory

The requirements for sampling are summarized by the Nyquist sampling theorem.[1] Let $x_c(t)$ be a band-limited signal with $X_c(j\Omega) = 0$ for $|\Omega| \geq \Omega_N$. Then $x_c(t)$ is uniquely determined by its samples, $x[n] = x_c(nT)$, if $\Omega_S = 2\pi/T \geq 2\Omega_N$. The frequency Ω_N is referred to as the *Nyquist frequency*, and the frequency $2\Omega_N$ is referred to as the *Nyquist rate*. This theory is significant because it states that as long as a continuous-time signal is band-limited and sampled at least twice as fast as the highest frequency, then it can be exactly reproduced by the sampled sequence.

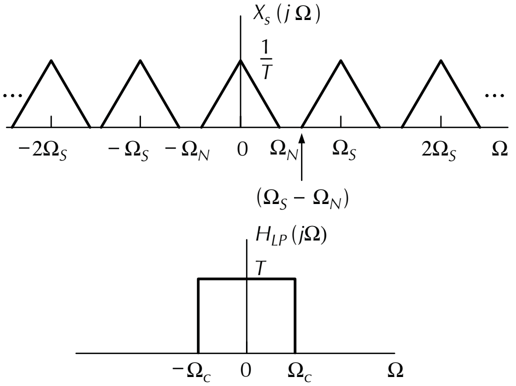


Figure 31-12. The spectrum replicas and the ideal low-pass filter that will remove the copies except for the desired baseband spectrum.

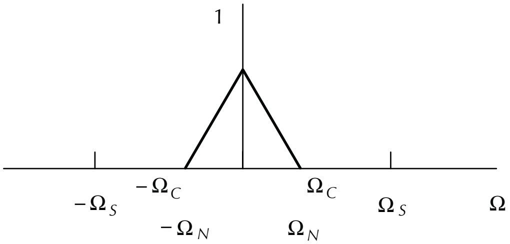


Figure 31-13. The final result of reconstructing the analog signal from the sampled signal.

The sampling analysis can be extended to the frequency response of the discrete time sequence, $x[n]$, by using the relationships $x[n] = x_c(nT)$ and

$$X(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} x[n]e^{-j\omega n}$$

The result is that

$$X(e^{j\omega}) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_c\left(j\left(\frac{\omega}{T} - \frac{2\pi k}{T}\right)\right) \qquad (31\text{-}20)$$

$X(e^{j\omega})$ is a frequency-scaled version of the continuous-time frequency response, $X_s(j\Omega)$, with the frequency scale specified by $\omega = \Omega T$. This scaling can also be thought of as normalizing the frequency axis by the sample rate so that frequency components that occurred at the sample rate now occur at 2π. Because the time axis has been normalized by the sampling period T, the frequency axis can be thought of as being normalized by the sampling rate $1/T$.

31.6.4 Quantization

The discussion up to this point has been on how to quantify the effects of periodically sampling a continuous-time signal to create a discrete-time version of the signal. As shown in Fig. 31-9, there is a second step—namely, mapping the infinite-resolution discrete-time signal into a finite precision representation (i.e., some number of bits per sample) that can be manipulated in a computer. This second step is known as *quantization*. The quantization process takes the sample from the continuous-to-discrete conversion and finds the closest corresponding finite precision value and represents this level with a bit pattern. This bit pattern code for the sample value is usually a binary two's-complement code so that the sample can be used directly in arithmetic operations without the need to convert to another numerical format (which takes some number of instructions on a DSP processor to perform). In essence, the continuous-time signal must be both quantized in time (i.e., sampled), and then quantized in amplitude.

The quantization process is denoted mathematically as

$x[n] = Q(x[n])$

where,

$Q(\bullet)$ is the nonlinear quantization operation,

$x[n]$ is the infinite precision sample value.

Quantization is nonlinear because it does not satisfy Eq. 31-7—i.e., the quantization of the sum of two values is not the same as the sum of the quantized values due to how the nearest finite precision value is generated for the infinite-precision value.

To properly quantize a signal, it is required to know the expected range of the signal— i.e., the maximum and minimum signal values. Assuming the signal amplitude is symmetric, the most positive value can be denoted as X_M. The signal then ranges from $+X_M$ to $-X_M$ for a total range of $2X_M$. Quantizing the signal to B bits will decompose the signal into 2^B different values. Each value represents $2X_M/2^B$ in amplitude and is represented as the step size $\Delta = 2X_M 2^{-B} = X_M 2^{-(B-1)}$. As a simplified example of the quantization process, assume that a signal will be quantized into eight different values which can be conveniently represented as a 3-bit value. Fig. 31-14 shows one method of how an input signal, $x[n]$, can be converted into a 3-bit quantized value, $Q(x[n])$. In this figure, values of the input signal between $-\Delta/2$ and $\Delta/2$ are given the value 0. Input signal values between $\Delta/2$ and $3/2$ are represented by their average value Δ, and so forth. The eight output values range from -4Δ to 3Δ for input signals between

$-9\Delta/2$ and $7\Delta/2$. Values larger than $7\Delta/2$ are set to 3Δ and values smaller than $-9\Delta/2$ are set to -4Δ—i.e., the numbers saturate at the maximum and minimum values, respectively.

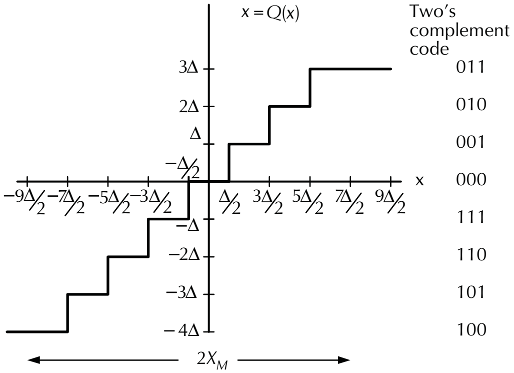


Figure 31-14. The quantization of an input signal, *x*, into *Q*(*x*).

The step size, Δ, has an impact on the resulting quality of the quantization. If Δ is large, fewer bits will be required for each sample to represent the range, $2X_M$, but there will be more quantization errors. If Δ is small, more bits will be required for each sample, although there will be less quantization error. Normally, the system design process determines the value of X_M and the number of bits required in the converter, *B*. If X_M is chosen too large, then the step size, Δ, will be large and the resulting quantization error will be large. If X_M is chosen too small, then the step size, Δ, will be small, but the signal may clip the A/D converter if the actual range of the signal is larger than X_M.

This loss of information during quantization can be modeled as noise signal that is added to the signal as shown in Fig. 31-15. The amount of quantization noise determines the overall quality of the signal. In the audio realm, it is common to sample with 24 bits of resolution on the A/D converter. Assuming a ±15 V swing of an analog signal, the granularity of the digitized signal is $30\text{V}/2^{24}$, which comes to 1.78 μV.

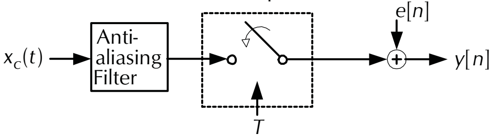


Figure 31-15. The sampling process of Fig. 31-9 with the addition of an antialiasing filter and modeling the quantization process as an additive noise signal.

With certain assumptions about the signal, such as the peak value being about four times the rms signal value, it can be shown that the signal to noise ratio (SNR) of the A/D converter is approximately 6 dB per bit.[1] Each additional bit in the A/D converter will contribute 6 dB to the SNR. A large SNR is usually desirable, but that must be balanced with overall system requirements, system cost, and possibly other noise issues inherent in a design that would reduce the value of having a high-quality A/D converter in the system. The dynamic range of a signal can be defined as the range of the signal levels over which the SNR exceeds a minimum acceptable SNR.

There are cost-effective A/D converters that can shape the quantization noise and produce a high-quality signal. Sigma-Delta converters, or noise-shaping converters, use an oversampling technique to reduce the amount of quantization noise in the signal by spreading the fixed quantization noise over a bandwidth much larger than the signal band.[5] The technique of oversampling and noise shaping allows the use of relatively imprecise analog circuits to perform high-resolution conversion. Most digital audio products on the market use these types of converters.

31.6.5 Sample Rate Selection

The sampling rate, *1/T*, plays an important role in determining the bandwidth of the digitized signal. If the analog signal is not sampled often enough, then high-frequency information will be lost. At the other extreme, if the signal is sampled too often, there may be more information than is needed for the application, causing unnecessary computation and adding unnecessary expense to the system.

In audio applications it is common to have a sampling frequency of 48 kHz = 48,000 Hz, which yields a sampling period of 1/48,000 = 20.83 μs. Using a sample rate of 48 kHz is why, in many product data sheets, the amount of delay that can be added to a signal is an integer multiple of 20.83 μs.

The choice of which sample rate to use depends on the application and the desired system cost. High-quality audio processing would require a high sample rate while low bandwidth telephony applications require a much lower sample rate. A table of common applications and their sample rate and bandwidths are shown in Table 31-3. As shown in the sampling process, the maximum bandwidth will always be less than ½ the sampling frequency. In practice, the antialiasing filter will have some roll-off and will band limit the signal to less than ½ the sample rate. This bandlimiting will further reduce the bandwidth, so the final bandwidth of the audio signal will be a function of the filters implemented in the specific A/D and the sample rate of the system.

Table 31-3. Common Sample Rates Found in Typical Applications and the Practical Bandwidths Realized at Each Sample Rate

Application	Sample Rate	Bandwidth
Telephony applications	8 kHz	3.5 kHz
Videoconferencing	16 kHz	7 kHz
FM radio	32 kHz	15 kHz
CD audio	44.1 kHz	20 kHz
Professional audio	48 kHz	22 kHz
Future audio	96 kHz	45 kHz

31.7 Algorithm Development

Once a signal is digitized, the next step in a DSP system is to process the signal. The system designer will begin the design process with some goal in mind and will use the algorithm development phase to develop the necessary steps (i.e., the algorithm) for achieving the goal.

The design cycle for a DSP system generally has three distinct phases as shown in Fig. 31-16: an abstract algorithm conceptualization phase, in which various mathematical algorithms and systems are explored; an algorithm development phase, where the algorithms are tested on large amounts of data; and a system implementation phase, where specific hardware is used to realize the system.

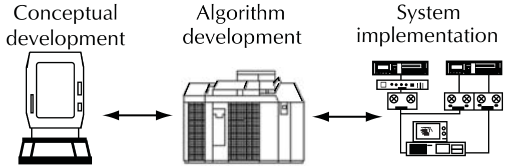


Figure 31-16. The three phases of DSP application development.

Traditionally, the three phases in the DSP design cycle have been performed by three entirely different groups of engineers using three entirely different classes of tools, although this process has converged as development tools have improved. The algorithm conceptualization phase is most often performed by researchers in a laboratory environment using highly interactive, graphically oriented DSP simulation and analysis tools. In this phase, the researcher begins with the concept of what to accomplish and creates the simulation environment that will enable changes and reformulations of the approach to the problem. No consideration is given, at this point, to the computational performance issues. The focus is on proof of concept issues—proving that the approach can solve the problem (or a temporarily simplified version of the problem).

In the algorithm development phase, the algorithms are fine-tuned by applying them to large databases of signals, often using high-speed workstations to achieve the required throughput. During this step it is often necessary to refine the high-level conceptualization in order to address issues that arose while running data through the system. Simulations are characterized by having many probes on the algorithm to show intermediate signal values, states, and any other useful information to aid in troubleshooting both the algorithm and the implementation of the simulation.

Once the simulation performs as desired, the next step is to create a real-time implementation of the simulation. The purpose of the real-time implementation is to better simulate the final target product, to begin to understand what the real-time memory and computational requirements will be, and to run real-time data through the system. There is no substitute for running real-time data through the system because real-time data typically exhibits characteristics that were either not anticipated or have unintended consequences in the simulation environment. Real-time data is generally more stressful to an algorithm than simulated, or non-real-time, data.

Often, with the introduction of real-time data, it may be necessary to go to the conceptual level again and further refine the algorithm.

Although advanced development tools and high-speed processors have blurred the distinction between simulation and real-time implementation, the goal of the real-time implementation is to "squeeze as much algorithm as possible" into the target processor (or processors). Squeezing more into the target processor is a desirable goal because it is usually much less expensive to use a single signal processor than to use multiple processors.

31.8 Digital Signal Processors

Programmable digital signal processors are microprocessors with particular features suited to performing arithmetic operations such as multiplication and addition very efficiently.[3,4] Traditionally, these enhancements have improved the performance of the processor at the expense of ease of programmability.

A typical microprocessor will have an arithmetic and logic unit for performing arithmetic operations, a memory space, I/O pins, and possible other peripherals such as serial ports and timers. A DSP processor will often have fewer peripherals, but will include a hardware multiplier, often a high-speed internal memory space, more memory addressing modes, an instruction cache and a pipeline, and even a separation of the program and data memory spaces to help speed program execution. The hardware multiplier allows the DSP processor to perform a multiplication in a single clock cycle while microprocessors typically take multiple clock cycles to perform this task. With clock cycles easily exceeding 100 MHz, up to 100 million multiples can occur every second. At this rate, 2083 multiplies can occur in the time span required to collect one sample of data at a 48 kHz sample rate (100 M/48,000).

A high-speed internal memory bank can be used to speed the access to the data and/or program memory space. By making the memory high speed, the memory can be accessed twice within a single clock cycle, allowing the processor to run at maximum performance. This means that proper use of internal memory enables more processing to take place within a given speed processor when compared to using external memory.

The instruction cache is also used to keep the processor running more efficiently because it stores recently used instructions in a special place in the processor where they can be accessed quickly, such as when looping program instructions over signal data.

The pipeline is a sequential set of steps that allow the processor to fetch an instruction from memory, decode the instruction, and execute the instruction. By running these subsystems in parallel, it is possible for the processor to be executing one instruction while it is decoding the next one and fetching the instruction after that. This streamlines the execution of instructions.

31.8.1 DSP Arithmetic

Programmable DSPs offer either fixed-point or floating-point arithmetic. Although floating-point processors are typically more expensive and offer less performance than fixed-point processors, VLSI hardware advances are minimizing the differences. The main advantage of a floating-point processor is the ability to be free of numerical scaling issues, simplifying the algorithm development and implementation process.

Most people naturally think in terms of fractions and decimal points, which are examples of floating-point numbers. Typically, floating-point DSPs can represent very large and very small numbers and use 32-bit (or longer) words composed of a 24-bit mantissa and an 8-bit exponent, which together provide a dynamic range from 2^{-127} to 2^{128}. This vast range in floating-point devices means that the system developer does not need to spend much time worrying about numerical issues such as overflow (a number too large to be represented) or underflow (a number too small to be represented). In a complicated system, there is enough to worry about without having to worry about numerical issues as well.

Fixed-point arithmetic is called fixed-point because it has a fixed decimal point position and because the numbers have an implicit scale, depending on the range that must be represented. This scale must be tracked by the programmer when performing arithmetic on fixed-point numbers. Most DSPs use the fixed-point 2s-complement format, in which a positive number is represented as a simple binary value and a negative value is represented by inverting all the bits of the corresponding positive value and then adding 1. Assuming a 16-bit word, there are $2^{16} = 65{,}536$ possible combinations or values that can be represented which allows the representation of numbers ranging from the largest positive number of $2^{15} - 1 = 32{,}767$ to the smallest negative (e.g., most negative) number of $-2^{15} = -32{,}768$.

There are many times when it is important to represent fractions in addition to integer numbers. To represent fractions, the implied position of the decimal point must be moved. When using 16-bit arithmetic to represent fractions only, with no integer component, a Q15 arithmetic format with an implied decimal point and 15 bits of fraction data to the right of the decimal point could be used. In this case, the largest number that can be represented is still $2^{15} - 1$, but now this number represents 32,767/32,768 = 0.999969482, and the smallest negative number is still -2^{15}, but this number represents –32,768/32,768 = –1. Using Q15 arithmetic, it is possible to represent numbers between 0.999969482 and –1. As another example, representing numbers that range between 16 and –16 would require Q11 arithmetic (4 bits before the implied decimal point). An implementation may use different implied decimal positions for different variables in a system.

Because of the smaller word size and simpler arithmetic operations when compared to floating-point processors, fixed-point DSPs typically use less silicon area than their floating-point counterparts, which translates into lower prices and less power consumption. The trade-off is that, due to the limited dynamic range and the rules of fixed-point arithmetic, an algorithm designer must play a more active role in the development of a fixed-point DSP system. The designer has to decide whether the given word width (typically 16 or 24 bits) will be interpreted as integers or fractions, apply scale factors if required, and protect against possible register overflows at potentially many different places in the code. Overflow occurs in two ways in a fixed-point DSP.[3] Either a register overflows when too many numbers are added to it or the program attempts to store N bits from the accumulator and the discarded bits are important. A complete solution to the overflow problem requires the system designer to be aware of the scaling of all the variables so that overflow is sufficiently unlikely. An underflow occurs if a number is smaller than the smallest number that can be represented. Floating-point arithmetic keeps track of the scaling automatically in order to simplify the programmer's job. The exponent keeps track of where the decimal point should be. Checking for overflow/underflow and preventing these conditions makes changing a DSP algorithm more difficult because, not only are algorithmic changes required, there are also numeric issues to contend with. Usually, once an implementation for a particular application has matured past the development stage, the code (which may have begun as floating-point code) may be ported to a fixed-point processor to allow the cost of the product to be reduced.

The dynamic range supported in a fixed-point processor is a function of the bit width of the processor's data registers. As with A/D conversion, each bit adds 6 dB to the SNR. A 24-bit DSP has 48 dB more dynamic range than a 16-bit DSP.

31.8.2 Implementation Issues

The implementation of an algorithm into a real system is often much more complicated than using a compiler to automatically optimize the code for maximum performance. Real-time systems have constraints such as limited memory, limited computational performance, and most importantly, need to handle the real-time data that is continuously sent from the A/D converter to the DSP and the real-time data that must be sent from the DSP back to the D/A converter. Interruptions in this real-time data are typically not acceptable because, for example, in an audio application, these interruptions will cause audible pops and clicks in the audio signal.

Real-time programming requires that all of the computation required to produce the output signal must happen within the amount of time it takes to acquire the input signal from the A/D converter. In other words, each time an input sample is acquired, an output sample must be produced. If the processing takes too long to produce the output, then, at some point, incoming data from the A/D will not be able to be processed, and input samples will be lost. As an example, assume a system samples at 48 kHz and performs parametric equalization on a signal. Assuming that each band of parametric equalization requires 5 multiplies and 4 adds, which can be implemented in 9 clock cycles, then a 100 MHz DSP has 2083 instructions that can be executed in the time between samples. These instructions would allow a maximum of 231 bands of parametric equalization (2083/9 = 231). Now, realistically, the system is performing other tasks such as collecting data from the A/D converter, sending data to the D/A converter, handling overhead from calling subroutines and returning from subroutines, and is possibly responding to interrupts from other subsystems. So the actual number of bands of equalization could be significantly less than the theoretical maximum of 231 bands.

DSPs will have a fixed amount of internal memory and a fixed amount of external memory that can addressed. Depending on the system to be designed, it can be advantageous to minimize the amount of external memory that is required in a system because that can lead to reduced parts costs, reduced manufacturing expense, and higher reliability. However, there is usually a trade-off between computational requirements and memory usage. Often, it is possible to trade memory space for increased computational power and vice versa. A simple example of this would be the creation of a sine wave. The DSP can either compute the samples of a sine wave, or look-up the values in a table. Either method will produce the appropriate sine wave, but the former will require less memory and more CPU while the latter will require more memory and less CPU. The system designer usually makes a conscious decision regarding which trade-off is more important.

31.8.3 System Delay

Depending on the application, one of the most important issues in an implementation is the amount of delay or latency that is introduced into the system by the sampling and processing. Fig. 31-17 shows the typical digital system. The analog signal comes into the A/D converter that digitizes and quantizes the signal. Once digitized, the signal is typically stored in some data buffers or arrays of data. The data buffers could be one sample long or could be longer depending on whether the algorithm operates on a sample-by-sample basis or requires a buffer of data to perform its processing. The system buffers are usually configured in a ping-pong fashion so that while one buffer is filling up with new data from the A/D, the other is being emptied by the DSP as it pulls data from the buffer to process the data.

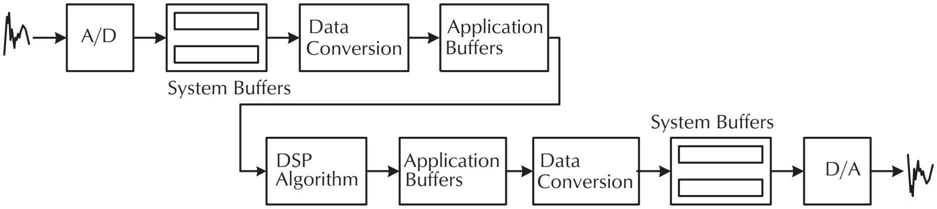


Figure 31-17. A block diagram of the typical DSP system.

Following the system buffer may be a data conversion block that converts the data from a fixed-point integer format provided by the A/D to either some other fixed-point format or a floating-point processor, depending on the DSP and the numerical issues. Following this, there may be some application buffers that store buffers of data to give the DSP some flexibility in how much time it takes to process a single block of data. The application buffers can be viewed as a rubber band that allows the DSP to use more time for some frames of data and less time for other frames of data. As long as the average amount of time required to process a buffer of data is less than the amount of time required to acquire that buffer of data, the DSP will make real-time. If the amount of time required to process a buffer takes longer than the time to acquire the buffer, then the system will be unable to process all buffers and will have to drop buffers because there will not be any processing time left over to collect the next buffer from the A/D converter. In this case the system will not make real time and the missing buffers will produce audible pops in an audio signal. The application buffers can be used to compensate for some frames that may require more processing (more CPU time) than others. By providing more frames over which to average the computation, the DSP will more likely make real time. Of course, if the DSP cannot perform the required amount of computation on average during the time that a buffer of data is acquired, then averaging over more and more frames will not help. The system will eventually miss real time and have to drop samples.

After the application buffers, the DSP algorithm performs the operations that are desired and then passes the data to possibly another set of application buffers that in turn can be converted from the numerical format of the DSP to the format required by the D/A converter. Finally the data will be sent to the D/A converter and converted back into an analog signal.

An accounting of the delay of the system should include all delays beginning when the analog signal comes in contact with the A/D converter to when the analog signal leaves the D/A converter. Table 31-4 shows the potential delays in each of the blocks of Fig. 31-17. For this exercise, it is assumed that a frame of data consists of N samples, where $N \geq 1$. Each frame of delay adds $N \bullet 1/T$ seconds of delay to the system. For example, a delay of 16 samples at 48 kHz corresponds to 16/48,000 = 333.3 µs.

Table 31-4. A Summary of Delay Issues in a Typical DSP System

Block	Delay	Description
A/D	From 1 to 16 samples	Most A/D converters have some amount of delay built in due to the processing that is done. Oversampling A/Ds in particular have more delay than other types of A/Ds.
System Buffers	Adds at least 1 frame of delay	In the ping-pong buffer scheme, the system is always processing the last frame of data while the A/D is supplying the data from the next frame of data.
Data conversion	Possibly none	The conversion of the data format may be lumped with the algorithm processing delay.
Application buffers	Adds M-1 frames of delay for M buffers	Generalizing the ping-pong buffer scheme to M buffers, the system is always processing the oldest buffer, which is M-1 buffers behind the most recent buffer.
DSP algorithm	Variable, although usually at least 1 frame	There are two primary ways a DSP algorithm adds delay. One is processing delay and the other is algorithmic delay. Processing delay occurs because the processor is not infinitely fast, so it takes some amount of time to perform all of the computation. If the DSP has no extra CPU cycles after performing the computation, then the processing time adds a full frame of delay to the system. If it takes more than a frame of delay to perform the computation, then the system will not make real time. The algorithmic delay comes from any requirement to use data from future frames of data (i.e., buffer the data) in order to make decisions about the current frames of data and other delays inherent in the algorithm process.
D/A	From 1 to 16 samples	As with the A/D converter there is some delay associated with converting a digital signal into an analog signal. Current converters typically have no more than 16 samples of delay.

Further complicating the delay measurements is the possible requirement of sending information to an external system. This could be in the form of sending a bitstream to a remote decoder, receiving a bitstream from a remote encoder, and also any error detection and/or correction on a bitstream that may be required.

31.8.4 Choosing a DSP

The choice of which DSP to use for a particular application depends on a collection of factors including:

- **Cost.** DSPs range in price from several dollars to hundreds of dollars. Low-cost DSP processors are typically 16-bit fixed-point devices with limited amounts of internal memory and few peripherals. Low-cost DSPs are typically suited for extremely high volume applications, where the exact capabilities required, and no more, are built into the chip.

  High-cost DSPs typically are newer processors that have a great deal of internal memory or other architectural features including floating-point arithmetic and high speed communication ports.

- **Computational Power: MHz, MIPs, MFLOPs.** Computational power is measured in several different ways including processor speed (MHz), millions of instructions per second (MIPS), and millions of floating-point operations per second (MFLOPS). The computational power of a processor is usually directly related to cost. An MIP means that one million instructions can be executed per second. The instructions that can be executed could include memory loads and stores or perhaps arithmetic operations. An MFLOP means one million floating-point operations can be executed per second. A floating-point operation includes multiplies and/or adds. Often the architecture of the DSP allows the DSP to execute two (or more) floating-point operations per instruction. In this case the MFLOPs would be twice (or more) the MIPs rating of the processor.

  Higher-speed processors allow the user to pack more features into a DSP product, but with a higher cost.

Power Consumption. Depending on the application, low power may be important for long battery life or low heat dissipation. DSPs will have a power rating and, often, a watt/MIP rating to estimate power consumption.

Architecture. Different manufacturers' DSPs have different features and trade-offs. Some processors may allow extremely high-speed computational rates but at the expense of being difficult to program. Some may offer ease of multiprocessing, multiple arithmetic processors, or other features.

Arithmetic Precision. The use of floating-point arithmetic simplifies arithmetic operations. Fixed-point processors often have lower cost but often require additional instructions to maintain the level of numerical accuracy that is often required. The final production volume of the end product often dictates whether the added development time is worth the cost savings.

Peripherals. Certain features of processors such as the ability to share processor resources among linked processors or access to external memory/devices can have a significant impact on which processor to use for a particular application. Integrated timers, serial ports, and other features can reduce the number of additional parts required in a design.

Code Development. The amount of code already developed for a particular processor family may dictate the choice of processors. Real-time code development takes significant time and the investment can be substantial. The ability to reuse existing code is a significant time saver in getting products to market.

Development Tools. The development tools are critical to the timely implementation of an algorithm on a particular processor. If the tools are not available or are not functional, the development process will most likely be extended beyond any reasonable time estimate.

Third Party Support. DSP processor manufacturers have a network of companies that provide tools, algorithm implementations, and hardware solutions for particular problems. It is possible that some company has already implemented, and makes a living out of implementing, the type of solution that is required for a given application.

31.9 Programming a DSP

DSPs, like many other processors, are only useful if they can input and output data. The software system used to input and output data is called an I/O system. As shown in Fig. 31-17, a DSP application program typically processes an input stream of data to produce some output data. The processing of this data is performed under the direction of the application program, which usually includes one or more algorithms programmed on the DSP. The DSP application program consists of acquiring the input stream data, using the algorithms to process the data, and then outputting the processed data to the output data stream. An example of this is a speech data compression system where the input stream is a data stream representing uncompressed speech. The output stream, in this case, is the compressed speech data and the application consists of getting the uncompressed input speech data, compressing the data, and then sending the compressed data to the output stream.

One of the most important factors that a DSP I/O system must address is the idea of real-time. An extremely important aspect of these real time A/D and D/A systems is that the samples must be produced and consumed at a fixed rate in order for the system to work in real-time. Although an A/D or D/A converter is a common example of a real-time device, other devices not directly related to real-time data acquisition can also have real time constraints. This is particularly true if they are being used to supply, collect, or transfer real-time information from devices such as disk drives and interprocessor communication links. In the speech compression example, the output stream might be connected to a modem that would transmit the compressed speech to another DSP system that would uncompress the speech. The I/O system should be designed to interface to these devices (or any other) as well.

Another important aspect of a real-time I/O system is the amount of delay imposed from input to output. For instance, when DSPs are used for in-room reinforcement or two-way speech communication (i.e., telecommunications), the delay must be minimized. If the DSP system causes a noticeable delay, the conversation would be awkward and the system would be considered unacceptable. Therefore, the DSP I/O system should be capable of minimizing I/O delay to a reasonable value.

Programming a DSP is usually accomplished in a combination of C and assembly languages. The C code provides a portable implementation that can potentially be run on multiple different platforms. Assembly language allows for a more computationally efficient implementation at the expense of increased development time and decreased portability. By starting in C, the developer can incrementally optimize the implementation by benchmarking which subroutines are taking the most time, optimizing these routines, and then finding the next subroutine to optimize.

The typical C code shell for implementing a DSP algorithm is shown in Fig. 31-18. Here, the C code allocates some buffer memory to store signal data, opens an I/O signal stream, and then gets data, processes the data, and then sends the data to the output stream. The input and output streams typically have lower level device drivers for talking directly to the A/D and D/A converters, respectively.

```
#include <stdio.h>
#include <aspi_io.h>
#include <malloc.h>

#define LEN 800

void main(argc,argp)
char **argp;
int argc;
{
   SIG_Stream input, output;
   SIG_Attrs sig_attrs;
   BUF_Buffer buffer;

   buffer = BUF_create(SEG_DRAM,LEN,0);

   input = SIG_open(argp[1],SIG_READ,buffer,0);

   SIG_getattrs(input,&sig_attrs);

        output =
            SIG_open(argp[2],SIG_WRITE,buffer,&sig_attrs);

  while (SIG_get(input,buffer))
                  {
       /* data processing of buffer */
       my_DSP_algorithm(buffer);

       SIG_put(output,buffer);
                  }
       return(0);
}
```

Figure 31-18. An example C program for collecting data from an A/D using an input signal stream created with SIG_open and sending data to the D/A using the output signal stream and processing the data with the function my_DSP_algorithm().

31.10 Conclusion

This chapter has introduced the fundamentals of DSP from a theoretical perspective (signal and system theory), and a practical perspective. The concepts of real-time systems, data acquisition, and digital signal processors have been introduced. DSP is a large and encompassing subject and the interested reader is encouraged to learn more through the exhaustive treatment given to this material in the references.[1,2]

References

1. Alan V. Oppenheim, Ronald W. Schafer, with John R. Buck, *Discrete-Time Signal Processing*, 2nd Edition, Upper Saddle River, NJ:Prentice Hall, 1999.
2. James H. McClellan, Ronald W. Schafer, Mark Yoder, *DSP First: A Multimedia Approach*, Upper Saddle River, NJ:Prentice Hall 1999.
3. Edward A. Lee, "Programmable DSP Architectures: Part I", *IEEE ASSP Magazine*, Vol. 5, no. 4, October 1988, pp. 4–19.
4. Edward A. Lee, "Programmable DSP Architectures: Part II", *IEEE ASSP Magazine*, Vol. 6, no. 1, January 1989, pp. 4–14.
5. Pervez M. Aziz, Henrik V. Sorensen, and Jan Van Der Spiegel, "An Overview of Sigma-Delta Converters," *IEEE Signal Processing Magazine*, Vol. 13, no. 1, January 1996, pp. 61–84.
6. Alan V. Oppenheim, A. S. Willsky, with S. H. Nawab, *Signals and Systems*, 2nd Edition, Upper Saddle River, NJ:Prentice Hall, 1997.

Chapter 32

Grounding and Interfacing

by Bill Whitlock

32.1 Introduction 1181
32.2 Basic Electronics 1181
  32.2.1 Circuit Theory 1181
  32.2.2 ac Circuits 1182
  32.2.3 Capacitance, Inductance, and Impedance 1182
  32.2.4 Single Wires 1183
  32.2.5 Cables and Transmission Lines 1184
32.3 Electronics of Interfaces 1185
  32.3.1 Balanced and Unbalanced Interfaces 1185
  32.3.2 Voltage Dividers and Impedance Matching 1186
  32.3.3 Line Drivers and Cable Capacitance 1187
  32.3.4 Capacitive Coupling and Shielding 1187
  32.3.5 Inductive Coupling and Shielding 1188
32.4 Grounding 1189
  32.4.1 Earth Grounding 1190
  32.4.2 Fault or Safety Grounding 1190
  32.4.3 Signal Grounding and EMC 1192
  32.4.4 Grounding and System Noise 1193
    32.4.4.1 Power Line Noise 1193
    32.4.4.2 Parasitic Capacitances and Leakage Current Noise 1194
    32.4.4.3 Parasitic Transformers and Inter-Outlet Ground Voltage Noise 1194
    32.4.4.4 Ground Loops 1194
32.5 Interface Problems in Systems 1195
  32.5.1 Degraded Common-Mode Rejection 1196
  32.5.2 The Pin 1 Problem 1198
  32.5.3 Balanced Cable Issues 1200
  32.5.4 Coupling in Unbalanced Cables 1202
  32.5.5 Bandwidth and RF Interference 1203
32.6 Solving Real-World System Problems 1203
  32.6.1 Noise Perspective 1203
  32.6.2 Troubleshooting 1204
    32.6.2.1 Observations, Clues, and Diagrams 1204
    32.6.2.2 The Ground Dummy Procedure 1204
32.6.2.2.1 For Unbalanced Interfaces 1205
32.6.2.2.2 For Balanced Interfaces 1206
  32.6.3 Solving Interface Problems 1207
    32.6.3.1 Ground Isolators 1207
    32.6.3.2 Multiple Grounding 1207
    32.6.3.3 Unbalanced to Balanced and Vice Versa 1208

    - 32.6.3.4 RF Interference 1210
    - 32.6.3.5 Signal Quality 1212
    - 32.6.3.6 Tips for Balanced Interfaces 1212
    - 32.6.3.7 Tips for Unbalanced Interfaces 1213
- 32.7 Alternative Treatments and Pseudoscience 1214
  - 32.7.1 Grounding from Technical to Bizarre 1214
  - 32.7.2 Power-Line Isolation, Filtering, and Balancing 1214
  - 32.7.3 Surge Protection 1216
  - 32.7.4 Exotic Audio Cables 1216
- References 1217
- Suggested Reading 1219
- Trademarks 1219

32.1 Introduction

Many audio professionals think of system grounding as a black art. How many times have you heard someone say that a cable is picking up noise, presumably from the air like a radio receiver? Even equipment manufacturers often don't have a clue what's really going on when there's a problem. The most basic rules of physics are routinely overlooked, ignored, or forgotten. As a result, myth and misinformation have become epidemic! This chapter is intended to enable sound engineers to understand and either avoid or solve real-world noise problems. The electronic system engineering joke that cables are sources of potential trouble connecting two other sources of potential trouble contains more truth than humor. Because equipment ground connections have profound effects on noise coupling at signal interfaces, we must appreciate how interfaces actually work as well as when, why, and how equipment is grounded. Although the subject can't be reduced to just a few simple rules, it doesn't involve rocket science or complex math either.

For convenience in this chapter, we'll use the term *noise* to mean to signal artifacts that originate from sources *external* to the signal path. This includes hum, buzz, clicks, or pops originating from the power line and interference originating from radio-frequency devices. A predictable amount of white noise is inherent in all electronic devices and must be expected. This random noise, heard as hiss, will also limit the usable dynamic range of any audio system, but this is not the subject of this chapter!

Any signal accumulates noise as it flows through the equipment and cables in a system. Once it contaminates a signal, noise is essentially impossible to remove without altering or degrading the signal. Therefore, noise and interference must be prevented along the entire signal path. It might seem trivial to transfer signal from the output of one audio device to the input of another but, in terms of noise and interference, signal interfaces are truly the danger zone! Let's start with some basic electronics that apply to interfaces.

32.2 Basic Electronics

Fields can exert invisible forces on objects within them. In electronics, we're concerned with electric and magnetic fields. Almost everyone has seen a demonstration of iron filings sprinkled on paper used to visualize the magnetic field between the north and south poles of a small magnet. A similar electric field exists between two points having a constant voltage difference between them. Fields like these, which neither move nor change in intensity, are called *static* fields.

If a field, either magnetic or electric, moves in space or fluctuates in intensity, the other kind of field will be generated. In other words, a changing electric field will set up a changing magnetic field or a changing magnetic field will set up a changing electric field. This interrelationship gives rise to *electromagnetic waves*, in which energy is alternately exchanged between electric and magnetic fields as they travel through space at the speed of light.

Everything physical is made of atoms whose outermost components are electrons. An electron carries a negative electric charge and is the smallest quantity of electricity that can exist. Some materials, called *conductors* and most commonly metals, allow their outer electrons to move freely from atom to atom. Other materials, called *insulators* and most commonly air, plastic, or glass, are highly resistant to such movement. This movement of electrons is called current flow. Current will flow only in a complete circuit consisting of a connected source and load. **Regardless of how complex the path becomes, all current leaving a source must return to it!**

32.2.1 Circuit Theory

An electric potential or *voltage*, sometimes called *emf* for electromotive force, is required to cause current flow. It is commonly denoted *E* (from emf) in equations and its unit of measure is the *volt*, abbreviated *V*. The resulting rate of current flow is commonly denoted *I* (from intensity) in equations and its unit of measure is the ampere, abbreviated *A*. How much current will flow for a given applied voltage is determined by circuit resistance. Resistance is denoted *R* in equations and its unit of measure is the ohm, symbolized Ω.

Ohm's Law defines the quantitative relationship between basic units of voltage, current, and resistance:

$$E = I \times R$$

which can be rearranged as

$$R = \frac{E}{I}$$

$$I = \frac{E}{R}$$

For example, a voltage *E* of 12 V applied across a resistance *R* of 6 Ω will cause a current flow *I* of 2 A.

Circuit elements may be connected in series, parallel, or combinations of both, Figs. 32-1 and 32-2.

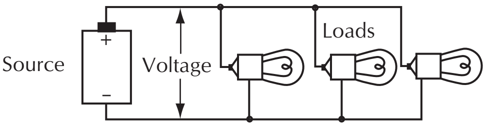


Figure 32-1. The voltage is the same across all elements in a parallel circuit.

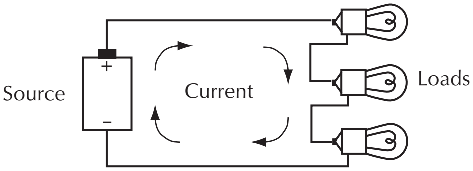


Figure 32-2. The current is the same through all elements in a series circuit.

Although the resistance of wires that interconnect circuit elements is generally assumed to be negligible, we will discuss this later.

In a parallel circuit, the total source current is the sum of the currents through each circuit element. The highest current will flow in the lowest resistance, according to Ohm's Law. The equivalent single resistance seen by the source is always lower than the lowest resistance element and is calculated as

$$R_{EQ} = \frac{1}{\frac{1}{R1} + \frac{1}{R2} + \frac{1}{R} \ldots + \frac{1}{n}} \tag{32-1}$$

In a series circuit, the total source voltage is the sum of the voltages across each circuit element. The highest voltage will appear across the highest resistance, according to Ohm's Law. The equivalent single resistance seen by the source is always higher than the highest resistance element and is calculated as

$$R_{EQ} = R1 + R2 + R3 \ldots + Rn \tag{32-2}$$

Voltages or currents whose value (magnitude) and direction (polarity) are steady over time are generally referred to as *dc*. A battery is a good example of a dc voltage source.

32.2.2 ac Circuits

A voltage or current that changes value and direction over time is generally referred to as *ac*. Consider the voltage at an ordinary 120 V, 60 Hz ac receptacle.

Since it varies over time according to a mathematical sine function, it is called a *sine wave*. Figure 32-3 shows how it would appear on an oscilloscope where time is the horizontal scale and instantaneous voltage is the vertical scale with zero in the center. The instantaneous voltage swings between peak voltages of +170 V and −170 V. A cycle is a complete range of voltage or current values that repeat themselves periodically (in this case every 16.67 ms). Phase divides each cycle into 360° and is used mainly to describe instantaneous relationships between two or more ac waveforms. Frequency indicates how many cycles occur per second of time. Frequency is usually denoted *f* in equations, and its unit of measure is the hertz, abbreviated Hz. Audio signals rarely consist of a single sine wave. Most often they are complex waveforms consisting of many simultaneous sine waves of various amplitudes and frequencies in the 20 Hz to 20,000 Hz (20 kHz) range.

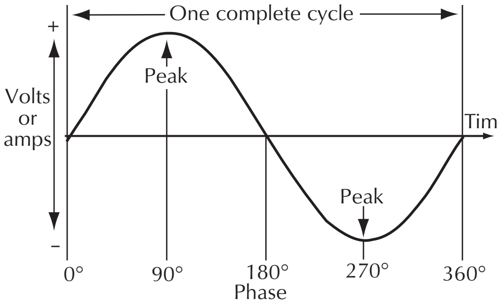


Figure 32-3. Sine wave as displayed on an oscilloscope.

32.2.3 Capacitance, Inductance, and Impedance

An electrostatic field exists between any two conductors having a voltage difference between them. Capacitance is the property that tends to oppose any change in the strength or charge of the field. In general, capacitance is increased by larger conductor surface areas and smaller spacing between them. Electronic components expressly designed to have high capacitance are called *capacitors*. Capacitance is denoted *C* in equations and its unit of measure is the Farad, abbreviated F. It's very important to remember that unintentional or parasitic capacitances exist virtually everywhere. As we will see, these parasitic capacitances can be particularly significant in cables and transformers!

Current must flow in a capacitor to change its voltage. Higher current is required to change the voltage rapidly and no current will flow if the voltage is held constant. Since capacitors must be alternately charged and discharged in ac circuits, they exhibit an apparent ac resistance called *capacitive reactance*. Capacitive reactance is inversely proportional to both capacitance and frequency since an increase in either causes an increase in current, corresponding to a decrease in reactance.

$$X_C = \frac{1}{2\pi f C} \qquad (32\text{-}3)$$

where,
X_C is capacitive reactance in ohms,
f is frequency in hertz,
C is capacitance in farads.

In general, capacitors behave as open circuits at dc and gradually become short circuits, passing more and more current, as frequency increases.

As shown in Figure 32-4, a magnetic field exists around any conductor carrying current at right angles to the axis of flow. The strength of the field is directly proportional to current. The direction, or polarity, of the magnetic field depends on the direction of current flow. Inductance is the property that tends to oppose any change in the strength or polarity of the field. Note that the fields around the upper and lower conductors have opposite polarity. The fields inside the loop point in the same direction, concentrating the field and increasing inductance. An electronic component called an *inductor* (or *choke*) is most often made of a wire coil with many turns to further increase inductance. Inductance is denoted L in equations and its unit of measure is the henry, abbreviated H. Again, remember that unintentional or parasitic inductances are important, especially in wires!

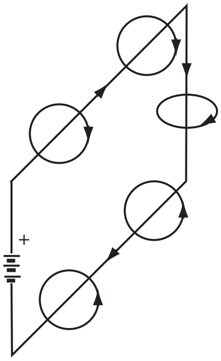


Figure 32-4. Magnetic field surrounding conductor.

If we abruptly apply a dc voltage to an inductor, a magnetic field is generated within the wire and moves outward as current begins to flow. But, in accordance with the law of induction, the rising field strength will induce a voltage, called *back emf*, in the wire which works to oppose current flow. The faster the field increases its strength, the more back emf will be induced to oppose current flow. The net result is to slow the buildup of current as it approaches its final value, which is limited by the applied voltage and circuit resistance. In ac circuits, for a constant applied voltage, this slowing reduces current flow as frequency increases because less time is available each cycle for current to rise. This apparent increase in ac resistance is called *inductive reactance*. Inductive reactance increases in direct proportion to both inductance and frequency.

$$X_L = 2\pi f L \qquad (32\text{-}4)$$

where,
X_L is inductive reactance in ohms,
f is frequency in hertz,
L is inductance in henrys.

In summary, inductors behave as short circuits at dc and gradually become open circuits, passing less and less current, as frequency increases.

Impedance is the combined effect of both resistance and reactance for circuits that contain resistance, capacitance, and inductance, which is the case with virtually all real-world circuits. Impedance is represented by the letter Z and is measured in ohms. Impedance can be substituted for R in the Ohm's Law equations. Impedance is a more general term than either resistance or reactance and, for ac circuits is the functional equivalent of resistance.

32.2.4 Single Wires

The electrical properties of wire are often overlooked. Consider a 10 ft length of #12 AWG solid copper wire.

1. The resistance of a wire is directly proportional to its length, inversely proportional to its diameter, and depends strongly on the material. From standard wire tables, we find the dc resistance of #12 AWG annealed copper wire is 1.59 Ω/1000 ft or 0.0159 Ω for a 10 ft length. At frequencies below about 500 Hz, this resistance largely sets the impedance.
2. The inductance of a straight wire is nearly independent of its diameter but is directly proportional to its length. From the formula for the inductance of a straight round wire,[1] we find its inductance is 4.8 µH. As shown in Fig. 32-5, this causes a rise in impedance beginning at about 500 Hz, reaching 30 Ω at 1 MHz (AM radio). Replacing the wire with a massive ½ inch diameter copper rod would reduce impedance only slightly to 23 Ω.

3. Electromagnetic waves travel through space or air at the speed of light. The physical distance traveled by a wave during one cycle is called *wavelength*. The equation is

$$M = \frac{984}{f} \tag{32-5}$$

where,
M is wavelength in feet,
f is frequency in MHz.

For 1 MHz AM radio, 100 MHz FM radio, and 2 GHz cell phone signals, wavelengths are about 1000 ft, 10 ft, and 6 inches, respectively.

4. Any wire will behave as an antenna at frequencies where its physical length is a quarter-wavelength or multiples thereof. This is responsible for the impedance peaks and dips seen at 25 MHz intervals in Fig. 32-5.

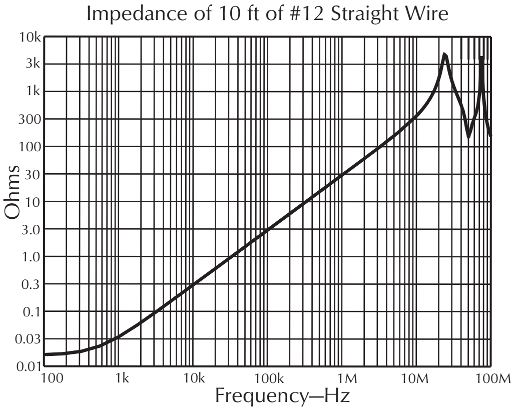


Figure 32-5. Wire is low-impedance current path only at low frequencies.

32.2.5 Cables and Transmission Lines

A cable consists of two or more conductors that are kept in close proximity over their length. Cables, such as those for ac power and loudspeakers, are generally used to convey power to a load. In a pair of such conductors, because the same current flows to and from the load in opposite directions, the magnetic fields have the same intensity but are of opposite polarity as shown in Fig. 32-6. In theory, there would be zero external field, and zero net inductance, if the two conductors could occupy the same space. The cancellation of round trip inductance due to magnetic coupling varies with cable construction, with typical values of 50% for zip cord, 70% for a twisted pair, and 100% for coaxial construction.

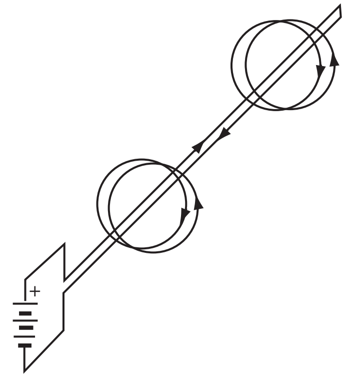


Figure 32-6. Cancellation of field in circuit pair.

At very high frequencies, a cable exhibits very different characteristics than it does at, say, 60 Hz power frequencies. This is caused by the finite speed, called *propagation velocity*, at which electrical energy travels in wires. It is about 70% of the speed of light for typical cables making wavelengths in cable correspondingly shorter. A cable is called electrically short when its physical length is under 10% of a wavelength at the highest frequency of interest. Wavelength at 60 Hz for typical cable is about 2200 miles (mi), making any power cable less than 220 mi long electrically short. Likewise, the wavelength at 20 kHz for typical cable is about 34,500 ft, making any audio cable less than about 3500 ft long electrically short. Essentially identical instantaneous voltage and current exists at all points on an electrically short cable and its signal coupling behavior can be represented by lumped resistance, capacitance, and magnetically coupled inductance as shown in Fig. 32-7. Its equivalent circuit can then be analyzed by normal network theory.

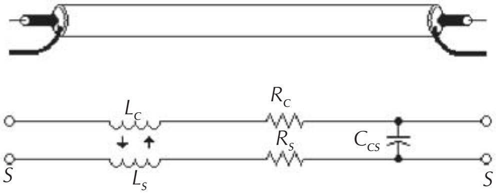


Figure 32-7. Lumped-circuit model of electrically short coaxial cable.

When a cable is longer than 10% of a wavelength, signals must be considered to propagate as electromagnetic waves and the cable can properly be called a *transmission line*. This includes typical cables longer than 7 ft for 10 MHz video, 8 inch for 100 MHz FM radio, and 0.8 inch for 1000 MHz CATV signals. Significantly different instantaneous voltages exist along the length of a transmission line. For all practical purposes, its electrical equivalent is a distributed circuit consisting of a large number of small inductors and resistors in series and capacitors in parallel. If an electrical impulse were applied to one end of an infinitely long cable, it would appear to have a purely resistive impedance. This *characteristic impedance* of the cable is a result of its inductance and capacitance per unit length, which is determined by its physical construction. Theoretically, the electrical impulse or wave would ripple down the infinite length of the cable forever. But actual transmission lines always have a far end. If the far end is left open or shorted, none of the wave's energy can be absorbed and it will reflect back toward the source. However, if the far end of the line is terminated with a resistor of the same value as the line's characteristic impedance, the wave energy will be completely absorbed. To the wave, the termination appears to be simply more cable. A properly terminated transmission line is often said to be matched. Generally, impedances of both the driving source and the receiving load are matched to the characteristic impedance of the line. In a mismatched line, the interaction between outgoing and reflected waves causes a phenomenon called *standing waves*. A measurement called *standing-wave ratio* (SWR) indicates mismatch, with an SWR of 1.00 meaning a perfect match.

32.3 Electronics of Interfaces

32.3.1 Balanced and Unbalanced Interfaces

An *interface* is a signal transport subsystem consisting of three components: a driver (one device's output), a line (interconnecting cable), and a receiver (another device's input). These components are connected to form a complete circuit for signal current, which requires a line having two signal conductors. The impedances of the signal conductors, usually with respect to ground, are what determine whether an interface is balanced or unbalanced. A concise definition of a balanced circuit is:

> *A balanced circuit is a two-conductor circuit in which both conductors and all circuits connected to them have the same impedance with respect to ground and to all other conductors. The purpose of balancing is to make the noise pickup equal in both conductors, in which case it will be a common-mode signal that can be made to cancel out in the load.*[2]

The use of balanced interfaces is an extremely potent technique to prevent noise coupling into signal circuits. It is so powerful that many systems, including telephone systems, use it in place of shielding as the main noise reduction technique![3]

Theoretically, a balanced interface can reject any interference, whether due to ground voltage differences, magnetic fields, or capacitive fields, as long as it produces identical voltages on each of the signal lines and the resulting peak voltages don't exceed the capabilities of the receiver.

A simplified balanced interface is shown in Fig. 32-8. Any voltage that appears on both inputs, since it is common to the inputs, is called a *common-mode* voltage. A balanced receiver uses a differential device, either a specialized amplifier or a transformer, that inherently responds only to the difference in voltage between its inputs. By definition, such a device will reject—i.e., have no response to—common-mode voltages. The ratio of differential gain to common-mode gain of this device is its *common-mode rejection ratio*, or CMRR. It's usually expressed in dB, and higher numbers mean more rejection. Section 32.5.1 will describe how CMRR often degrades in real-world systems and how it has traditionally been measured in ways that have no relevance to real-world system performance.

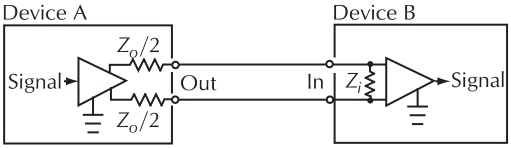


Figure 32-8. Basic balanced interface.

Two signal voltages have symmetry when they have equal magnitudes but opposite polarities. Symmetry of the desired signal has advantages, but they concern head room and crosstalk, not noise or interference rejection. The noise or interference rejection property is independent of the presence of a desired differential signal. Therefore, it can make no difference whether the desired signal exists entirely on one line, as a greater voltage on one line than the other, or as equal voltages on both of them. However, the symmetry myth is widespread. A typical example is: Each conductor is always equal in voltage but opposite in polarity to the other. The circuit that receives this signal in the mixer is called a *differential amplifier* and this opposing polarity of the conductors is essential for its operation. Like many others, it describes a balanced interface in terms of signal symmetry but never mentions impedances! Even the BBC test for output balance is actually a test for signal symmetry.[4] The idea that balanced interface is somehow defined by signal symmetry is simply wrong! It has apparently led some designers, mostly of exotic audiophile gear, to dispense with a differential amplifier input stage in their push-pull amplifiers. They simply amplify the (assumed) symmetrical input signals in two identical, ground-referenced amplifier chains. No mechanism exists to reject common-mode voltage (noise and interference) and it is actually amplified along with the signal, creating potentially serious problems. Rejection of common-mode voltages is the single most important function of a balanced receiver.

In an unbalanced circuit, one signal conductor is grounded (near-zero impedance) and the other has some higher impedance. As we will discuss in Section 32.5.4, the fact that not only signal but ground noise currents flow and cause voltage drops in the grounded conductor makes an unbalanced interface inherently susceptible to a variety of noise problems.

32.3.2 Voltage Dividers and Impedance Matching

Every driver has an internal impedance, measured in ohms, called its *output impedance*. Actual output impedance is important, as we discuss below, but often absent from equipment specifications. Sometimes, especially for consumer gear, the only impedance associated with an output is listed as recommended load impedance. While useful if listed in addition to output impedance, it is not what we need to know! A perfect driver would have a zero output impedance but, in practical circuit designs, it's neither possible nor necessary. Every receiver has an internal impedance, measured in ohms, called its *input impedance*. A perfect receiver would have an infinite input impedance but again, in practical circuit designs, it's neither possible nor necessary.

Figs. 32-8 and 32-9 illustrate ideal interfaces. The triangles represent ideal amplifiers having infinite impedance input—i.e., draw no current—and zero impedance output—i.e., deliver unlimited current—and the line conductors have no resistance, capacitance, or inductance. The signal voltage from the driver amplifier causes current flow through the driver output impedance(s) Z_o, the line, and receiver input impedance Z_i. Note that the output impedance of the balanced driver is split into two equal parts. Because current is the same in all parts of a series circuit and voltage drops are proportional to impedances, this circuit is called a *voltage divider*.

The goal of an interface is, with rare exception, to deliver maximum signal voltage from the output of one device to the input of another. Making Z_i much larger than Z_o assures that most of the signal voltage is delivered to the receiver and very little is lost in the driver. In typical devices, Z_o ranges from 30 Ω to 1 kΩ and Z_i ranges from 10 kΩ to 100 kΩ, which transfers 90–99.9% of the available—i.e., unloaded or open circuit—signal voltage.

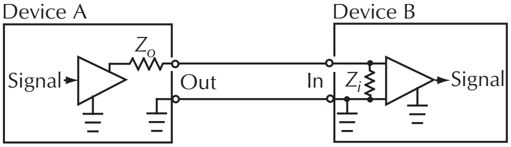


Figure 32-9. Basic unbalanced interface.

Matching is a term that often causes confusion. A little math and Ohm's Law will prove that when Z_o and Z_i are equal, maximum power is transferred from source to load, although half the signal voltage is lost. If transmission line effects apply, Z_o and Z_i must terminate or match the characteristic impedance of the line to prevent reflection artifacts. Although modern audio systems seldom use cables long enough for transmission line effects to apply or benefit from maximum power transfer, early telephone systems did both. Telephone systems began by using miles of existing open wire telegraph lines that, due to their wire size and spacing, had a characteristic impedance of 600 Ω. Since amplifiers didn't yet exist, the system was entirely passive and needed to transfer maximum power from one phone to another. Therefore, transformers, filters, and other components were designed for 600 Ω impedances to match the lines. These components were eventually incorporated into early sound reinforcement, radio, and recording systems. And the 600 Ω legacy still lives on, even though modern requirements for it are all but extinct.

Sometimes, instead of meaning equal, matching is used to mean optimizing some aspect of circuit performance. For example, the output transformer in a vacuum-tube power amplifier is used to optimize power output by converting or impedance matching the low-impedance loudspeaker to a higher impedance that suits the characteristics of the tubes. Similarly, the modern technique of making Z_i much larger than Z_o to transfer maximum voltage in signal interfaces is often referred to as *voltage matching*. It uses 10 kΩ or higher input impedances, called bridging because many inputs can be paralleled across the same output line with negligible drop in level. About 60 Ω has been suggested as the optimum Z_o for driving up to 2000 ft of typical shielded twisted pair cable in these balanced interfaces.[5]

32.3.3 Line Drivers and Cable Capacitance

A line driver and cable interact in two important ways. First, output impedance Z_o and the cable capacitance form a low-pass filter that will cause high-frequency roll-off. A typical capacitance for either unbalanced or balanced shielded audio cable might be about 50 pF/ft. If output impedance were 1 kΩ (not uncommon in unbalanced consumer gear), response at 20 kHz would be –0.5 dB for 50 ft, –1.5 dB for 100 ft, and –4 dB for 200 ft of cable. If the output impedance were 100 Ω (common in balanced pro gear), the effects would be negligible for the same cable lengths. Low-output impedance is especially important when cable runs are long. Also be aware that some exotic audio cables have extraordinarily high capacitance.

Second, cable capacitance requires additional high-frequency current from the driver. The current required to change the voltage on a capacitance is directly proportional to the rate of change or *slew rate* of the voltage. For a sine wave,

$$SR = 2\pi f V_p \tag{32-6}$$

where,

SR is slew rate in volts per second,

f is frequency in hertz,

V_p is peak voltage.

$$I = SR \times C \tag{32-7}$$

where,

I is current in A,

SR is slew rate in V/µs,

C is capacitance in µF.

For example, we have a cable with a slew rate of 1 V/µs at 20 kHz for a sine-wave of 8 V_p or 5.6 V_{rms}, which is also +17 dBu or +15 dBV. For a cable of 100 ft at 50 pF/ft, C would be 5000 pF or 0.005 µF. Therefore, peak currents of 5 mA are required to drive just the cable capacitance to +17 dBu at 20 kHz. Obviously, increasing level, frequency, cable capacitance, or cable length will increase the current required. Under the previous conditions, a cable of 1000 ft would require peak currents of 50 mA. Such peak currents may cause protective current limiting or clipping in the op-amps used in some line drivers. Since it occurs only at high levels and high frequencies, the audible effects may be subtle.

Of course, the load at the receiver also requires current. At a +17 dBu level, a normal 10 kΩ balanced input requires a peak current of only 0.8 mA. However, a 600 Ω termination at the input requires 13 mA. Matching 600 Ω sources and loads not only places a current burden on the driver but, because 6 dB (half) of signal voltage is lost, the driver must generate +23 dBu to deliver +17 dBu to the input. Unnecessary termination wastes driver current and unnecessary matching of source and load impedances wastes head room!

32.3.4 Capacitive Coupling and Shielding

Capacitances exist between any two conductive objects, even over a relatively large distance. As we mentioned earlier, the value of this capacitance depends on the surface areas of the objects and the distance. When there are ac voltage differences between the objects, these capacitances cause small but significant currents to flow from one object to another by means of the changing electric field (widely referred to as *electrostatic* fields although technically a misnomer since static means unchanging).

Strong electric fields radiate from any conductor operating at a high ac voltage and, in general, weaken rapidly with distance. Factors that increase coupling include increasing frequency, decreasing spacing of the wires, increasing length of their common run, increasing impedance of the victim circuit, and increasing distance from a ground plane. For some of these factors, there is a point of diminishing returns. For example, for parallel 22-gauge wires, there is no significant reduction in coupling for spacing over about 1 in.[6] Capacitive coupling originates from the voltage at the source. Therefore, coupling from a power circuit, for example, will exist whenever voltage is applied to the circuit regardless of whether load current is flowing.

Capacitive coupling can be prevented by placing electrically conductive material called a *shield* between the two circuits so that the electric field, and the resulting current flow, linking them is diverted. A shield is connected to a point in the circuit where the offending current will be harmlessly returned to its source, usually called *ground*—more about ground later. For example, capacitive coupling between a sensitive printed wiring board and nearby ac power wiring could be prevented by locating a grounded metal plate (shield) between them, by completely enclosing the board in a thin metal box, or by enclosing the ac power wiring in a thin metal box.

Similarly, as shown in Fig. 32-10, shielding can prevent capacitive coupling to or from signal conductors in a cable. Solid shields, such as conduit or overlapped foil, are said to have 100% coverage. Braided shields, because of the tiny holes, offer from 70% to 98% coverage. At very high frequencies, where the hole size becomes significant compared with interference wavelength, cables with combination foil/braid or multiple braided shields are sometimes used.

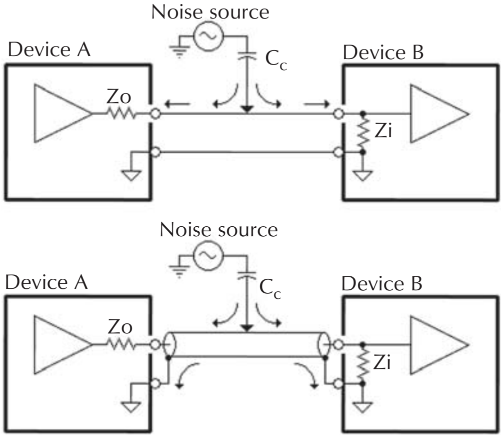


Figure 32-10. Capacitive noise coupling.

32.3.5 Inductive Coupling and Shielding

When any conductor cuts magnetic lines of force, in accordance with the law of induction, a voltage is induced in it. If an alternating current flows in the conductor, as shown at the left in Fig. 32-11, the magnetic field also alternates, varying in intensity and polarity. We can visualize the magnetic field, represented by the concentric circles, as expanding and collapsing periodically. Because the conductor at the right cuts the magnetic lines of force as they move across it, an ac voltage is induced over its length. This is the essential principle of a transformer. Therefore, current flowing in a wire in one circuit can induce a noise voltage in another wire in a different circuit. Because the magnetic field is developed only when current flows in the source circuit, noise coupling from an ac power circuit, for example, will exist only when load current actually flows.

If two identical conductors are exposed to identical ac magnetic fields, they will have identical voltages induced in them. If they are series connected as shown in Fig. 32-12, their identical induced voltages tend to cancel. In theory, there would be zero output if the two conductors could occupy the same space.

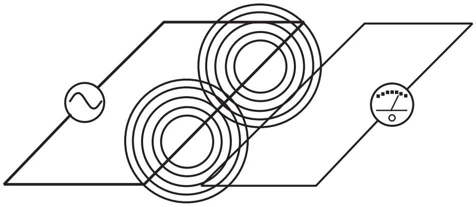


Figure 32-11. Inductive coupling between wires.

Magnetic fields become weaker rapidly as distance from the source increases, usually as the square of the distance. Therefore, cancellation depends critically on the two conductors being at precisely the same distance from the magnetic field source. Twisting essentially places each conductor at the same average distance from the source. So-called *star quad* cable uses four conductors with those opposing each other connected in parallel at each cable end. The effective magnetic center for each of these pairs is their center line and the two sets of pairs now have coincident center lines reducing the loop area to zero. Star quad cable has approximately 100 times (40 dB) better immunity to power-frequency magnetic fields than standard twisted pair. The shield of a coaxial cable also has an average location coincident with the center conductor. These construction techniques are widely used to reduce susceptibility of balanced signal cables to magnetic fields. In general, a smaller physical area inside the loop results in less magnetic radiation as well as less magnetic induction.

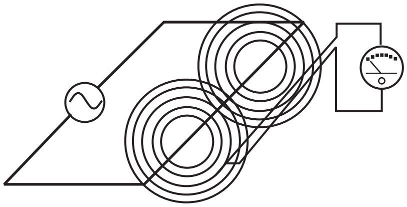


Figure 32-12. Coupling cancellation in loop.

Another way to reduce magnetic induction effects is shown in Fig. 32-13. If two conductors are oriented at a 90° (right) angle, the second doesn't cut the magnetic lines produced by the first and will have zero induced voltage. Therefore, cables crossing at right angles have minimum coupling and those running parallel have maximum coupling. The same principles also apply to circuit board traces and internal wiring of electronic equipment.

Magnetic circuits are similar to electric circuits. Magnetic lines of force always follow a closed path or circuit, from one magnetic pole to the opposite pole, always following the path of least resistance or highest

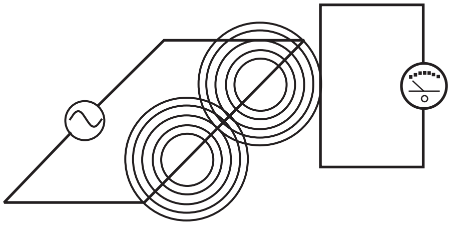


Figure 32-13. Zero coupling at right angles.

conductivity. The magnetic equivalent of electric current and conductivity are flux density and permeability. Hig-permeability materials have the ability to concentrate the magnetic force lines or flux. The permeability of air and other nonmagnetic materials such as aluminum, plastic, or wood is 1.00. The permeability of common ferromagnetic materials is about 400 for machine steel, up to 7000 for common 4% silicon transformer steel, and up to 100,000 for special nickel alloys. The permeability of magnetic materials varies with flux density. When magnetic fields become very intense, the material can become saturated, essentially losing its ability to offer an easy path for any additional flux lines. Higher permeability materials also tend to saturate at a lower flux density and to permanently lose their magnetic properties if mechanically stressed.

The basic strategy in magnetic shielding is to give the flux lines a much easier path to divert them around a sensitive conductor, circuit, or device. In general, this means that the shield must be a complete enclosure with a high magnetic permeability. The choice of the most effective shielding material depends on frequency. At low frequencies, below say 100 kHz, high-permeability magnetic materials are most effective. We can calculate how effective a conduit or cable shield will be at low frequencies:

$$SE = 20\log\left(1 + \frac{\mu t}{d}\right) \qquad (32\text{-}8)$$

where,

SE is shielding effect in dB,

μ is permeability of shield material,

t and d are the thickness and diameter (in the same units) of the conduit or shield.[7]

Thus, standard 1 inch EMT, made of mild steel with a low-frequency permeability of 300, will provide about 24 dB of magnetic shielding at low frequencies, but this will diminish to zero around 100 kHz. Fortunately, only low-frequency magnetic fields are generally a problem. In severe cases, nesting one magnetic shield inside another may be necessary.

Typical copper braid or aluminum foil cable shielding has little effect on magnetic fields at audio frequencies. If a shield is grounded at both ends, it behaves somewhat like a shorted turn to shield the inner conductors from magnetic fields.[8] Depending on the external impedance between the grounded ends of a cable shield, it may begin to become effective against magnetic fields somewhere in the 10 kHz to 100 kHz range. Box shields of aluminum or copper are widely used to enclose RF circuits because they impede magnetic fields through this eddy current action and are excellent shielding for electric fields as well. There is an excellent explanation of this high-frequency shielding in reference 9. However, copper or aluminum shielding is rarely an effective way to prevent noise coupling from audio-frequency magnetic fields.

32.4 Grounding

Historically, grounding became necessary for protection from lightning strokes and industrially-generated static electricity—i.e., belts in a flour mill. As utility power systems developed, grounding became standard practice to protect people and equipment. As electronics developed, the common return paths of various circuits were referred to as ground, regardless of whether or not they were eventually connected to earth. Thus, the very term *ground* has become vague, ambiguous, and often fanciful. Broadly, the purpose of grounding is to electrically interconnect conductive objects, such as equipment, in order to minimize voltage differences between them. An excellent general definition is that a *ground is simply a return path for current*, which will always return to its source. The path may be intentional or accidental—electrons don't care and don't read schematics![10]

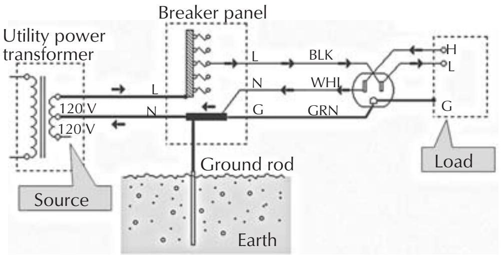
Grounding-related noise can be the most serious problem in any audio system. Common symptoms include hum, buzz, pops, clicks, and other noises. Because equipment manufacturers so often try to explain away these problems with the nebulous term *bad grounding*, most system installers and technicians feel that the entire subject is an incomprehensible black art. Adding to the confusion are contradictory rules proposed by various experts. Ironically, most universities teach very little about the real-world aspects of grounding. Graduates take with them the grounding fantasy that all grounds are equipotential—that is, have the same voltage. The fantasy certainly allows them to avoid complicated real-world interpretation of all those ground symbols on a schematic diagram, but the same fantasy can lead to noise disaster in their audio equipment and system designs.

Grounding has several important purposes and most often a single ground circuit serves, intentionally or accidentally, more than one purpose. We must understand how these ground circuits work and how noise can couple into signal circuits if we expect to control or eliminate noise in audio systems.

32.4.1 Earth Grounding

An *earth ground* is one actually connected to the earth via a low-impedance path. In general, earth grounds are necessary only to protect people from lightning. Before modern standards such as the National Electrical Code (NEC or just Code) were developed, lightning that struck a power line was often effectively routed directly into buildings, starting a fire or killing someone. Lightning strikes are the discharge of giant capacitors formed by the earth and clouds. Strikes involve millions of volts and tens of thousands of amperes, producing brief bursts of incredible power in the form of heat, light, and electromagnetic fields. Electrically, lightning is a high-frequency event, with most of its energy concentrated in frequencies over 300 kHz! That's why, as we discussed in Section 32.2.4, wiring to ground rods should be as short and free of sharp bends as possible. The most destructive effects of a strike can be avoided by simply giving the current an easy, low-impedance path to earth before it enters a building. Because overhead power lines are frequent targets of lightning, virtually all modern electric power is distributed on lines having one conductor that is connected to earth ground frequently along its length.

Fig. 32-14 shows how ac power is supplied through a three-wire split single-phase service to outlets on a typical 120 Vac branch circuit in a building. One of the service wires, which is often uninsulated, is the grounded neutral conductor. Note that both the white neutral and the green safety ground wires of each branch circuit are tied or bonded to each other and an earth ground rod (or its equivalent grounding electrode system) at the service entrance as required by Code. *This earth ground, along with those at neighboring buildings and at the utility poles, provide the easy paths for lightning to reach earth.*



Other phase of the incoming power , meter, main breaker, etc. omitted for clarity

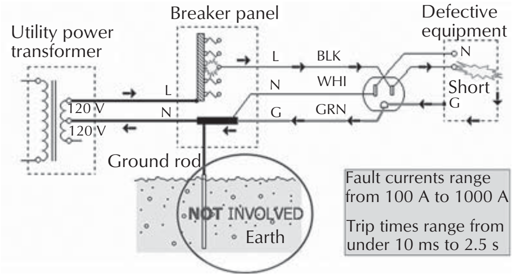
Figure 32-14. Simplified residential ac power from feeder to outlet.

Telephone, CATV, and satellite TV cables are also required to divert or arrest lightning energy before it enters a building. The telco-supplied gray box or NIU provides this protection for phone lines as x grounding blocks do for CATV and satellite dishes. NEC Articles 800, 810, and 820 describe requirements for telephone, satellite/TV antennas, and CATV, respectively. All protective ground connections should be made to the same ground rod used for the utility power, if the ground wire is 20 ft or less in length. If longer, separate ground rods must be used, and they must be bonded to the main utility power grounding electrode with a #6 AWG wire.[11] Otherwise, because of considerable soil resistance between separate ground rods, thousands of volts could exist between them when lightning events occur or downed power lines energize the signal lines. Without the bond such events could seriously damage a computer modem, for example, that straddles a computer grounded to one rod via its power cord and a telephone line protectively grounded to another.[12]

32.4.2 Fault or Safety Grounding

Any ac line powered device having conductive exposed parts (which includes signal connectors) can become a shock or electrocution hazard if it develops certain internal defects. Insulation is used in power transformers, switches, motors, and other internal parts to keep the electricity where it belongs. But, for various reasons, the insulation may fail and effectively connect live power to exposed metal. This kind of defect is called a *fault*. A washing machine, for example, could electrocute someone who happened to touch the machine and a water faucet (assumed grounded via buried metal pipes) at the same time.

NEC requires that 120 Vac power distribution in homes and buildings use a three-wire system as shown in Fig. 32-15. To prevent electrocution, most devices have a third wire connecting exposed metal to the safety ground pin of these outlets. The outlet safety ground is routed, through either the green wire or metallic conduit, to the neutral conductor and earth ground at the main breaker panel. The connection to neutral allows high fault current to flow, quickly tripping the circuit breaker, while the earth ground connection minimizes any voltage that might exist between equipment and other earth-grounded objects, such as water pipes, during the fault event. Power engineers refer to voltage differences created by these fault events as *step* or *touch* potentials. The neutral (white) and line (black) wires are part of the normal load circuit that connects the voltage source to the load. The green wire or conduit is intended to carry fault currents only.



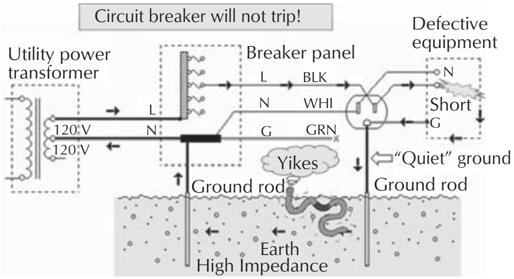
Fault currents range from 100 A to 1000 A

Trip times range from under 10 ms to 2.5 s

Figure 32-15. Fault protection is provided by safety ground to neutral bond.

NEC also requires safety grounding of wiring raceways and equipment cabinets, including rack cabinets. Per Article 250-95, safety grounding wires, which may be bare or insulated, must have a minimum size of #14 copper for a 15 A or #12 copper for a 20 A branch circuit to assure rapid circuit breaker action. This grounding path must be bonded to the safety grounding system, not to building steel or a separate earth ground system! *Separate earth grounds cannot provide safety grounding*!! As shown in Fig. 32-16, soil resistance is far too high to guarantee tripping of a circuit breaker under fault conditions.[13] With safety grounds in place, potentially deadly equipment faults simply cause high currents from power line hot to safety ground, quickly tripping circuit breakers and removing power from those branch circuits. Safety grounding in many residential and commercial buildings is provided through metal conduit, metallic J-boxes, and saddle-grounded or SG outlets. Technical or isolated grounding will be discussed in Section 32.7.

When trying to track down and correct system noise problems, it easy to assume that power outlets are wired correctly. Low-cost outlet testers, which generally cost less than $10.00, will find dangerous problems such as hot-neutral or hot-ground reversals and open connections. Because they check for correct voltages between the pins, and both neutral and ground are normally at 0 V, they cannot detect a neutral-ground reversal. This insidious wiring error can create nightmarish noise problems in an audio system. Finding the error by visual inspection of outlets is one possibility, but this could get labor intensive if the number of outlets is large. For large systems, and even those that can't be powered down, a sensitive, noncontact, clamp-on current probe can help identify the forks in the road when troubleshooting.[14] Code requires that neutral and safety ground be bonded only at the power service disconnecting means that is generally at the main breaker panel. Serious system noise problems can also occur when an extraneous neutral-to-ground connection exists elsewhere in the building wiring. A special test procedure can be used to determine this condition.[15]



Circuit breaker will not trip!

Figure 32-16. Fault protection is not provided by an earth ground connection!

NEVER, NEVER use devices such as three-prong-to two-prong ac plug adapters—a.k.a. ground lifters—to solve a noise problem! Such an adapter is intended to provide a safety ground (via the cover plate screw to a grounded saddle outlet and J-box) in cases where three-prong plugs must be connected to two-prong receptacles in pre-1960 buildings, Fig. 32-17.

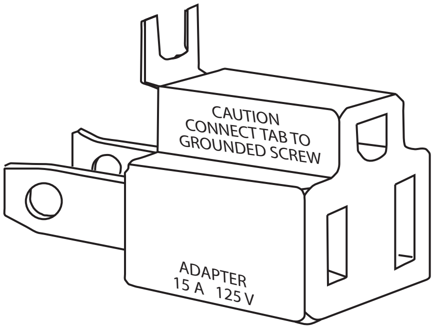


Figure 32-17. This is intended to provide a safety ground.

Consider two devices with grounding ac plugs that are connected by a signal cable. One device has a ground lifter on its plug and the other doesn't. If a fault occurs in the lifted device, the fault current flows through the signal cable to get to the grounded device. It's very likely that the cable will melt and catch fire!

Also consider that consumer audio and video equipment is responsible for about ten electrocutions every year in the United States. In a typical year, this equipment causes some 2000 residential fires that result in 100 civilian injuries, 20 deaths, and over $30 million in property losses, Fig.32-18[16,17]

Figure 32-18. Interconnect cables can carry lethal voltages throughout a system if just one ground lifted device fails.

Some small appliances, power tools, and consumer electronics are supplied with two-prong (ungrounded) ac plugs. Sometimes called *double insulated*, these devices are specially designed to meet strict UL and other requirements to remain safe even if one of their two insulation systems fails. Often there is a one-shot thermal cutoff switch inside the power transformer or motor windings to prevent overheating and subsequent insulation breakdown. *Only devices that carry a UL-listed label and originally supplied with ungrounded ac plugs should ever be operated without safety grounding*. Devices originally supplied with grounding three-prong plugs must always be operated with the safety ground properly connected!

32.4.3 Signal Grounding and EMC

EMC stands for electromagnetic compatibility, which is a field concerned with interference from electronic devices and their susceptibility to the interference created by other devices. As the world becomes increasingly wireless and digital, the general electromagnetic environment is becoming increasingly hostile. Engineers working in other disciplines, most notably information technology or IT—where signal/data frequencies are very high and narrowband—tend to minimize our difficulties in making audio systems robust against hostile electrical environments. In fact, high-quality audio systems are unique among electronic systems in two ways:

1. The signals cover a very broad, nearly 5 decade, range of frequencies.
2. The signals can require a very wide, currently over 120 dB, dynamic range.

Adding to the difficulty is the fact that ac power frequencies and their harmonics also fall within the system's working frequency range. As you might suspect, grounding plays a pivotal role in controlling both emissions and susceptibility in both electronic devices and systems. In general, the same principles and techniques that reduce emissions will also reduce susceptibility. Grounding schemes generally fall into one of three categories:

1. Single point or star grounding.
2. Multipoint or mesh grounding.
3. Frequency selective transitional or hybrid grounding.

At frequencies below about 1 MHz (which includes audio), virtually all experts agree that star grounding works best because system wiring is electrically short compared to the wavelengths involved. At these low frequencies, the dominant noise coupling problems arise from the simple lumped parameter behavior of wiring and electronic components. This includes the resistance and inductance of wires, the noise currents resulting from capacitances between utility power and system grounds, and magnetic and capacitive coupling effects.

On the other hand, at higher frequencies, system wiring can become electrically long and transmission line effects, such as standing waves and resonances, become the dominant problems. For example, because a 25 ft (7.5 m) audio cable is a quarter wavelength long at 10 MHz, it becomes an antenna. At frequencies of 100 MHz or higher, even a 12 in (30 cm) wire can no longer be considered low impedance path. To be effective at these frequencies, therefore, grounding schemes must emulate a flat metal sheet having extremely low inductance called a *ground plane*. In practice, this can usually only be approximated with a multipoint ground system using wires. The wire lengths between points must remain well under a quarter-wavelength so, as frequency increases, larger numbers of increasingly shorter wires must be used to create the mesh. Ultimately, only a real ground plane can produce low-impedance ground connections at very high frequencies. Even a ground plane is not a perfect or equipotential—i.e., zero volts at all points—ground. Because it has finite resistance, significant voltage differences can be developed between connection points to it.[18] Therefore, it should come as no surprise that IT and RF engineers prefer mesh grounding techniques while audio engineers prefer star grounding techniques.

At power and audio frequencies, a so-called *ground loop* allows noise and signal currents to mix in a common wire. Single-point grounding avoids this by steering signal currents and noise currents in independent paths. But at ultrasonic and radio frequencies, noise currents tend to bypass wires because they look like inductors and tend to flow instead in unintended paths consisting of parasitic capacitances. This makes star grounding essentially useless in controlling high-frequency interference in practical systems. Mesh grounding does a better job of controlling high-frequency interference, but since many ground loops are formed, low-frequency noise can easily contaminate signals. For audio systems, sometimes even inside audio equipment, there is clearly a conflict.

This conflict can be resolved by the hybrid grounding scheme. Capacitors can be used to create multiple high-frequency ground connections while allowing audio-frequency currents to take a path determined by the directly wired connection. Thus, the ground system behaves as a star system at low frequencies and a mesh system at high frequencies.[19] This technique of combining ground plane and star grounding is quite practical at the physical dimensions of a circuit board or an entire piece of equipment. At the system level the same conflict exists regarding grounding of audio cable shields. Ideally, at low frequencies, a shield should be grounded at one end only, but for maximum immunity to RF interference it should be grounded at both ends (and even intermediate points, if possible). This situation can be resolved by grounding one end directly and the other end through a small capacitor.[20] The shield grounding issue will be discussed further in Section 32.5.2.

32.4.4 Grounding and System Noise

Most real-world systems consist of at least two devices that are powered by utility ac power. These power line connections unavoidably cause significant currents to flow in ground conductors and signal interconnect cables throughout a system. Properly wired, fully Code-compliant premises ac wiring generates small ground voltage differences and leakage currents. They are harmless from a safety viewpoint but potentiality disastrous from a system noise viewpoint. Some engineers have a strong urge to reduce these unwanted voltage differences by shorting them out with a large conductor. The results are most often disappointing.[21] Other engineers think that system noise can be improved experimentally by simply finding a better or quieter ground. They hold a fanciful notion that noise current can somehow be skillfully directed to an earth ground, where it will disappear forever![22] In reality, since the earth has resistance just like any other conductor, earth ground connections are not at zero volts with respect to each other or any other mystical or absolute reference point.

32.4.4.1 Power Line Noise

The power line normally consists of a broad spectrum of harmonics and noise in addition to the pure 60 Hz sine wave voltage. The noise is created by power supplies in electronic equipment, fluorescent lights, light dimmers, and intermittent or sparking loads such as switches, relays, or brush-type motors (i.e., blenders, shavers, etc.). Fig. 32-19 shows how sudden changes in load current, zero to full in about a microsecond for a triac light dimmer in this case, generate bursts of high-frequency noise on the power line 120 times per second. Even an ordinary light switch will briefly arc internally as it is switched off and its contacts open, generating a single similar burst. This noise contains significant energy to at least 1 MHz that is launched into the power wiring. The wiring behaves like a complex set of misterminated transmission lines gone berserk, causing the energy to reflect back and forth throughout the premises wiring until it is eventually absorbed or radiated. Power line noise can couple into signal paths in several ways, usually depending on whether the equipment uses two-prong or three-prong (grounding) ac power connections.

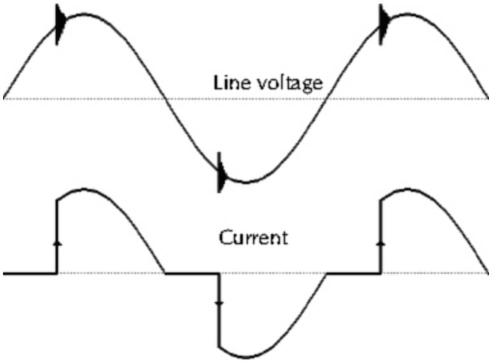


Figure 32-19. Light dimmer noise.

32.4.4.2 Parasitic Capacitances and Leakage Current Noise

In every ac-powered device, parasitic capacitances (never shown in schematic diagrams!) always exist between the power line and the internal circuit ground and/or chassis because of the unavoidable interwinding capacitances of power transformers and other line connected components. Especially if the device contains anything digital, there may also be intentional capacitances in the form of power line interference filters. These capacitances cause small but significant 60 Hz leakage currents to flow between power line and chassis or circuit ground in each device. Because the coupling is capacitive, current flow increases at higher noise frequencies. Fig. 32-20 shows the frequency spectrum of current flow in 3 nF of capacitance connected between line and safety ground at an ac outlet in a typical office.

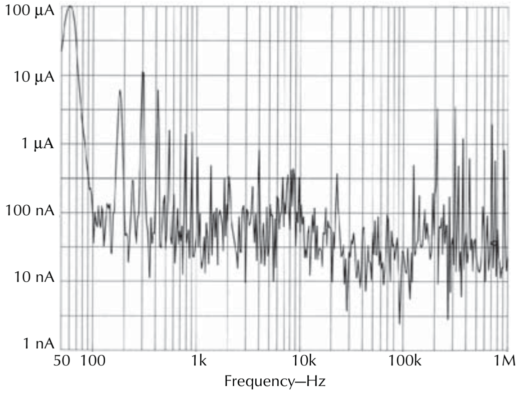


Figure 32-20. Typical leakage current from line to safety ground coupled via 3000 pF capacitance into a 75 Ω spectrum analyzer input.

This tiny current, although it poses no shock hazard, causes hum, buzz, pops, clicks, and other symptoms when it couples into the audio signal path. This capacitive coupling favors higher frequencies, making buzz a more common symptom than pure hum. We must accept noisy leakage currents as a fact of life.

32.4.4.3 Parasitic Transformers and Inter-Outlet Ground Voltage Noise

Substantial voltages are magnetically induced in premises safety ground wiring when load current flows in the circuit conductors as shown in Fig. 32-21. The magnetic fields that surround the line and neutral conductors, which carry load current, magnetically induce a small voltage over the length of the safety ground conductor, effectively forming a parasitic transformer. The closer the safety ground conductor is to either the line or neutral conductor, the higher the induced voltage. Because, at any instant in time, line and neutral currents are equal but flow in opposite directions, there is a plane of zero magnetic field exactly midway between the line and neutral conductors as shown in Fig. 32-22. Therefore, Romex® and similar bonded cables generally generate significantly lower induced voltages than individual wires in conduit, where the relative positioning of the wires is uncontrolled.

The voltage induced in any transformer is directly proportional to the rate of change of load current in the circuit. With an ordinary phase-control light dimmer the peak voltages induced can become quite high. When the dimmer triggers current on 120 times per second, it switches on very quickly (a few microseconds) as shown in Fig. 32-23. Since the magnetic induction into safety ground favors high frequencies, noise coupling problems in a system will likely become most evident when a light dimmer is involved. The problems are usually worst at about half-brightness setting of the dimmer.

This parasitic transformer action generates small ground voltage differences, generally under 1 V, between ac outlets. The voltage differences tend to be higher between two outlets on different branch circuits, and higher still if a device on the branch circuit is also connected to a remote or alien ground such as a CATV feed, satellite dish, or an interbuilding tie line. *We must accept interoutlet ground noise voltage as a fact of life.*

32.4.4.4 Ground Loops

For our purposes, a ground loop is formed when a signal cable connects two pieces of equipment whose connections to the power line or other equipment causes a power-line-derived current to flow in the signal cable.

The first, and usually worst, kind of ground loop occurs between *grounded* devices—those with three-prong ac plugs. Current flow in signal cables, as shown in Fig. 32-24, can easily reach 100 mA or more.

The second kind of ground loop occurs between floating devices—those with two-prong ac plugs. Each pair of capacitances CF (for EMI filter) and CP (for power transformer parasitic) in the schematic form a capacitive voltage divider between line and neutral, causing some fraction of 120 Vac to appear between chassis and ground. For UL-listed ungrounded equipment, this leakage current must be under 0.75 mA (0.5 mA for office equipment). This small current can cause an unpleasant, but harmless, tingling sensation as

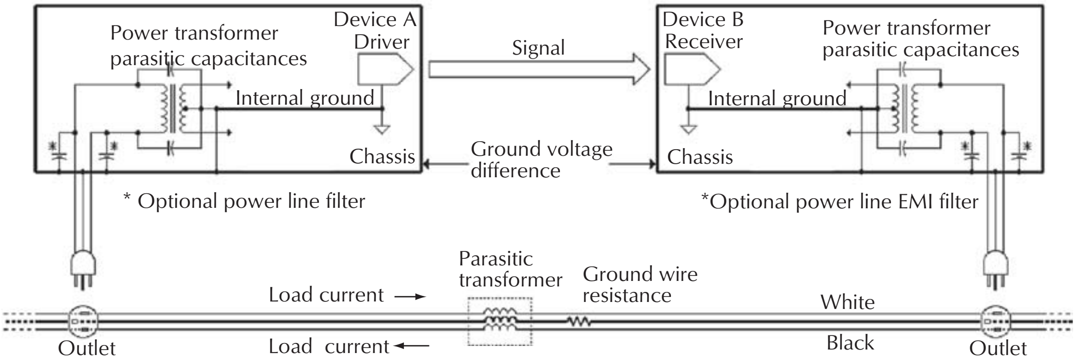

Power transformer
parasitic capacitances
Device A
Driver
Signal
Device B
Receiver
Power transformer
parasitic capacitances
Internal ground
Internal ground
Chassis
Ground voltage
difference
Chassis
* Optional power line filter
*Optional power line EMI filter
Parasitic
transformer
Ground wire
resistance
Load current
White
Load current
Black
Outlet
Outlet

Figure 32-21. Voltage difference is magnetically induced over length of safety-ground premises wiring.

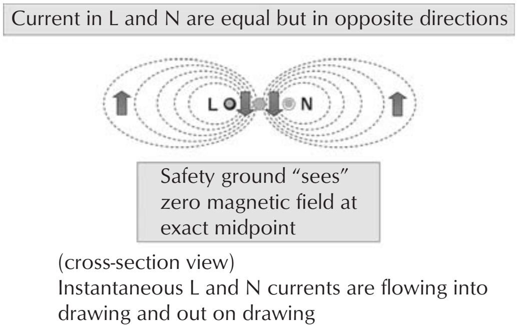

Current in L and N are equal but in opposite directions
L
N
Safety ground "sees"
zero magnetic field at
exact midpoint
(cross-section view)
Instantaneous L and N currents are flowing into
drawing and out on drawing

Figure 32-22. Magnetic fields surrounding line and neutral can induce voltage into safety ground.

it flows through a person's body. More relevant is the fact that these noisy leakage currents will flow in any wire connecting such a floating device to safety ground, or connecting two floating devices to each other as shown in Fig. 32-25.

32.5 Interface Problems in Systems

If properly designed balanced interfaces were used throughout an audio system, it would theoretically be noise-free. Until about 1970, equipment designs allowed real-world system to come very close to this ideal. But since then, balanced interfaces have fallen victim to two major design problems—and both can properly be blamed on equipment manufacturers. Even careful examination of manufacturers' specifications and data sheets will not reveal either problem—the devil is in the details. These problems are effectively concealed because the marketing departments of most manufacturers have succeeded in dumbing down their so-called specifications over the same time period.

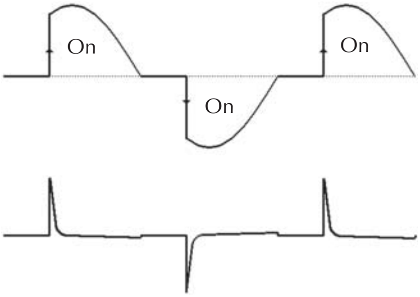


Figure 32-23. Lamp current (upper) versus induced voltage (lower) for phase-controlled dimmer.

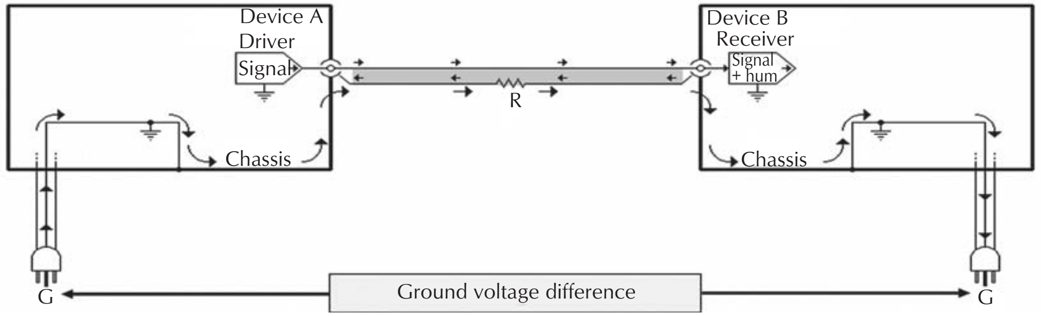

Device A
Driver
Signal
Device B
Receiver
Signal
+ hum
R
Chassis
Chassis
G
Ground voltage difference
G

Figure 32-24. For grounded equipment, interconnect cables complete a wired loop.

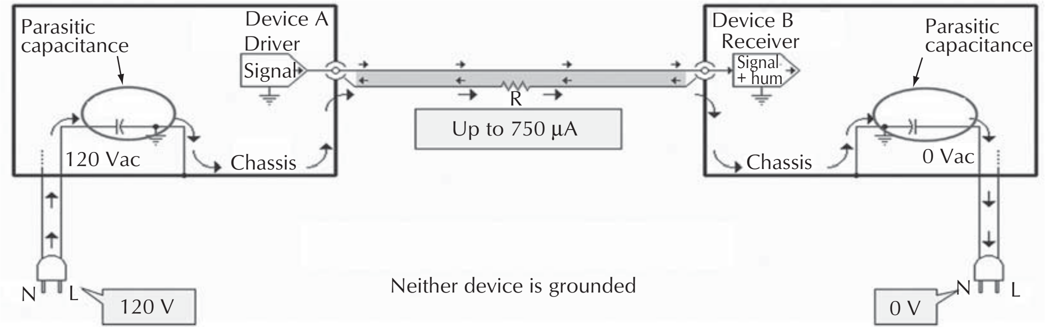

Parasitic capacitance
Device A
Driver
Signal
Device B
Receiver
Signal + hum
Parasitic capacitance
R
Up to 750 μA
120 Vac
Chassis
Chassis
0 Vac
N
L
120 V
Neither device is grounded
N
L
0 V

Figure 32-25. For ungrounded equipment, interconnect cables complete a capacitive loop.

First is degraded noise rejection, which appeared when solid-state differential amplifiers started replacing input transformers. Second is the pin 1 problem that appeared in large numbers when PC boards and plastic connectors replaced their metal counterparts. Both problems can be avoided through proper design, of course, but in this author's opinion, part of the problem is that the number of analog design engineers who truly understand the underlying issues is dwindling and engineering schools are steering most students into the digital future where analog issues are largely neglected. Other less serious problems with balanced interfaces are caused by balanced cable construction and choices of cable shield connections.

On the other hand, unbalanced interfaces have an intrinsic problem that effectively limits their use to only the most electrically benign environments. Of course, even this problem can be solved by adding external ground-isolation devices, but the best advice is to avoid them whenever possible in professional systems!

32.5.1 Degraded Common-Mode Rejection

Balanced interfaces have traditionally been the hallmark of professional sound equipment. In theory, systems comprised of such equipment are completely noise-free. However, an often overlooked fact is that the common-mode rejection of a complete signal interface does not depend solely on the receiver, but on how the receiver interacts with the driver and the line performing as a subsystem.

In the basic balanced interface of Fig. 32-26, the output impedances of the driver $Z_o/2$ and the input impedances of the receiver Z_{cm} effectively form the Wheatstone bridge shown in Fig. 32-27. If the bridge is not balanced or nulled, a portion of the ground noise V_{cm} will be converted to a differential signal on the line. This nulling of the common-mode voltage is critically dependent on the ratio matching of the pairs of driver/receiver common-mode impedances R_{cm} in the – and + circuit branches. The balancing or nulling is unaffected by impedance across the two lines, such as the signal input impedance Z_i in Fig. 32-28 or the signal output impedance of the driver. It is the common-mode impedances that matter!

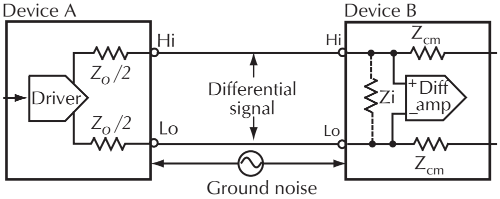

Device A
Device B
Hi
Hi
$Z_o/2$
$Z_o/2$
Driver
Differential signal
Lo
Lo
Z_{cm}
Z_{cm}
Zi
+ Diff _ amp
Ground noise

Figure 32-26. Simplified balanced interface.

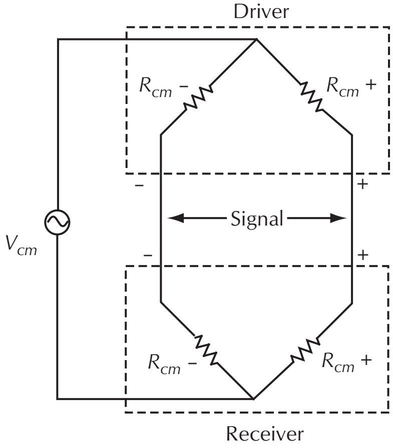

Driver
R_{cm} –
R_{cm} +
–
+
Signal
V_{cm}
–
+
R_{cm} –
R_{cm} +
Receiver

Figure 32-27. The balanced interface is a Wheatstone bridge.

The bridge is most sensitive to small fractional impedance changes in one of its arms when all arms have the same impedance.[23] It is least sensitive when the upper and lower arms have widely differing impedances—e.g., when upper arms are very low and lower arms are very high, or vice versa. Therefore, we can minimize the sensitivity of a balanced system (bridge) to impedance imbalances by making common-mode impedances very low at one end of the line and very high at the other. This condition is consistent with the requirements for voltage matching discussed in Section 32.3.2.

Most active line receivers, including the basic differential amplifier of Fig. 32-28, have common-mode input impedances in the 5 kΩ to 50 kΩ range, which is inadequate to maintain high CMRR with real-world sources. With common-mode input impedances of 5 kΩ, a source imbalance of only 1 Ω, which could arise from normal contact and wire resistance variations, can degrade CMRR by 50 dB. Under the same conditions, the CMRR of a good input transformer would be unaffected because of its 50 MΩ common-mode input impedances. Fig. 32-29 shows computed CMRR versus source imbalance for different receiver common-mode input impedances. Thermal noise and other limitations place a practical limit of about 130 dB on most actual CMRR measurements.

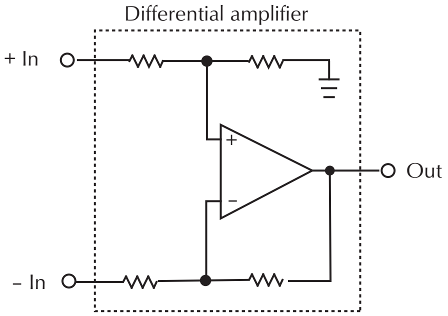


Figure 32-28. Basic differential amplifier.

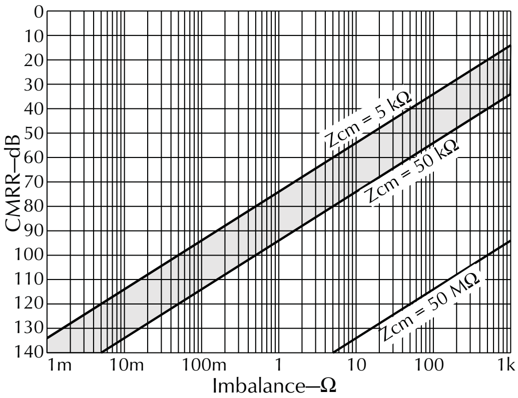


Figure 32-29. Noise rejection versus source impedance/imbalance.

How much imbalance is there in real-world signal sources? Internal resistors and capacitors determine the output impedance of a driver. In typical equipment, $Z_o/2$ may range from 25 to 300 Ω. Since the resistors are commonly ±5% tolerance and the coupling capacitors are ±20% at best, impedance imbalances up to about 20 Ω should be routinely expected. This defines a real-world source. In a previous paper, this author has examined balanced audio interfaces in some detail, including performance comparisons of various receiver types.[24] It was concluded that, regardless of their circuit topology, popular active receivers can have very poor CMRR when driven from such real-world sources. The poor performance of these receivers is a direct result of their low common-mode input impedances. If common-mode input impedances are raised to about 50 MΩ, 94 dB of ground noise rejection is attained from a completely unbalanced 1 kΩ source, which is typical of consumer outputs. When common-mode input impedances are sufficiently high, an input can be considered truly universal, suitable for any source—balanced or unbalanced. A receiver using either a good input transformer or the InGenius® integrated circuit[25] will routinely achieve 90–100 dB of CMRR and remain unaffected by typical real-world output imbalances.

The theory underlying balanced interfaces is widely misunderstood by audio equipment designers. Pervasive use of the simple differential amplifier as a balanced line receiver is evidence of this. And, as if this weren't bad enough, some have attempted to improve it. Measuring input X and Y input impedances of the simple differential amplifier individually leads some designers to alter its equal resistor values. However, as shown in Fig. 32-30, if the impedances are properly measured simultaneously, it becomes clear that nothing is wrong. The fix grossly unbalances the common-mode impedances, which destroys the interface CMRR for any real-world source. This and other misguided *improvements* completely ignore the importance of common-mode input impedances.

The same misconceptions have also led to some CMRR tests whose results give little or no indication of how the tested device will actually behave in a real-world system. Apparently, large numbers of designers test the CMRR of receivers with the inputs either shorted to each other or driven by a laboratory precision signal source. The test result is both unrealistic and misleading. Inputs rated at 80 dB of CMRR could easily deliver as little as 20 dB or 30 dB when used in a real system. Regarding their previous test, the

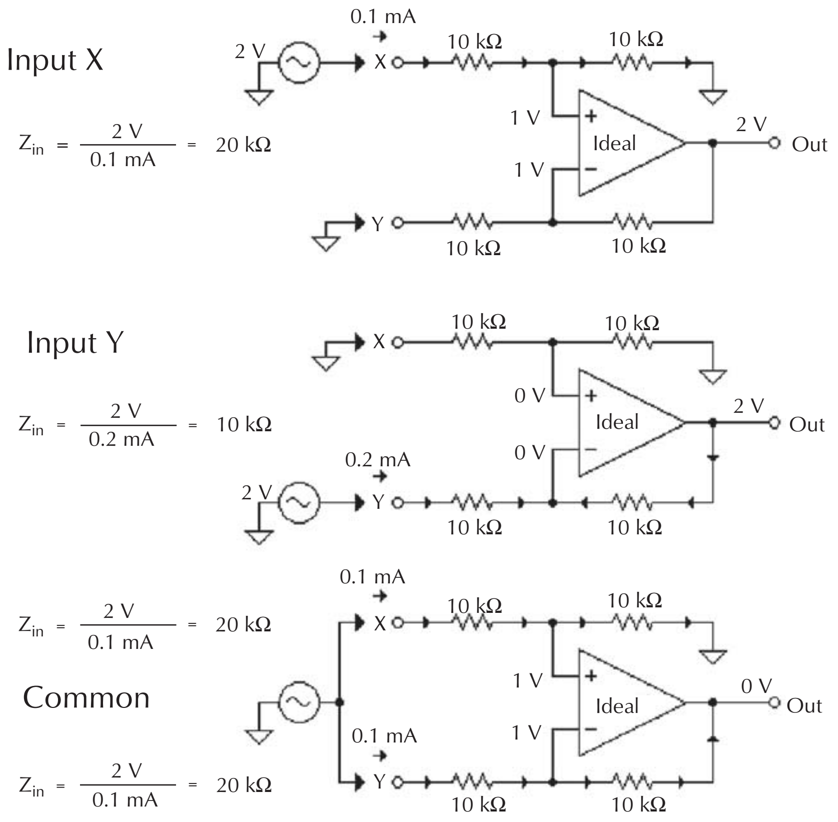

Input X
0.1 mA
2 V
X
10 kΩ
10 kΩ
$Z_{in} = \frac{2\ V}{0.1\ mA} = 20\ k\Omega$
1 V
Ideal
1 V
2 V
Out
Y
10 kΩ
10 kΩ
Input Y
X
10 kΩ
10 kΩ
$Z_{in} = \frac{2\ V}{0.2\ mA} = 10\ k\Omega$
0 V
Ideal
0 V
2 V
Out
0.2 mA
2 V
Y
10 kΩ
10 kΩ
0.1 mA
$Z_{in} = \frac{2\ V}{0.1\ mA} = 20\ k\Omega$
X
10 kΩ
10 kΩ
Common
1 V
Ideal
1 V
0 V
Out
0.1 mA
$Z_{in} = \frac{2\ V}{0.1\ mA} = 20\ k\Omega$
Y
10 kΩ
10 kΩ

Figure 32-30. Common-mode impedances apply to a voltage applied to both inputs.

IEC had recognized *that test is not an adequate assurance of the performance of certain electronically balanced amplifier input circuits*. The old method simply didn't account for the fact that source impedances are rarely perfectly balanced. To correct this, this author was instrumental in revising IEC Standard: 60268-3 Sound System Equipment – Part 3: Amplifiers. The new method, as shown in Fig. 32-31, uses typical ±10 Ω source impedance imbalances and clearly reveals the superiority of input transformers and some new active input stages that imitate them. The new standard was published August 30, 2000. The Audio Precision APx520 and APx525, introduced in 2008, are the first audio instruments to offer the new CMRR test.

32.5.2 The Pin 1 Problem

In his now famous paper in the 1995 *AES Journal*, Neil Muncy says:

> *This paper specifically addresses the problem of noise coupling into balanced line-level signal interfaces used in many professional applications, due to the unappreciated consequences of a popular and widespread audio equipment design practice which is virtually without precedent in any other field of electronic systems.*[30]

Common impedance coupling occurs whenever two currents flow in a shared or common impedance. A noise coupling problem is created when one of the currents is ground noise and the other is signal. The common impedance is usually a wire or circuit board trace having a very low impedance, usually well under an ohm. Unfortunately, common impedance coupling has been designed into audio equipment from many manufacturers. The noise current enters the equipment via a terminal, at a device input or output, to which the cable shield is connected via a mating connector. For XLR connectors, it's pin 1 (hence the name); for ¼ inch connectors, it's the sleeve; and for RCA/IHF connectors, it's the shell.

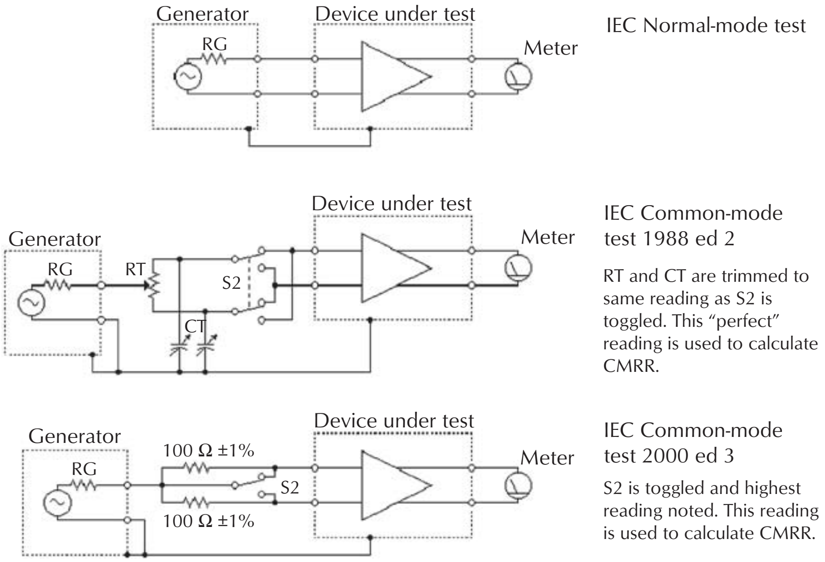

Generator
Device under test
RG
Meter
IEC Normal-mode test
Device under test
Generator
RG
RT
S2
CT
Meter
IEC Common-mode
test 1988 ed 2
RT and CT are trimmed to
same reading as S2 is
toggled. This "perfect"
reading is used to calculate
CMRR.
Device under test
Generator
RG
100 Ω ±1%
S2
100 Ω ±1%
Meter
IEC Common-mode
test 2000 ed 3
S2 is toggled and highest
reading noted. This reading
is used to calculate CMRR.

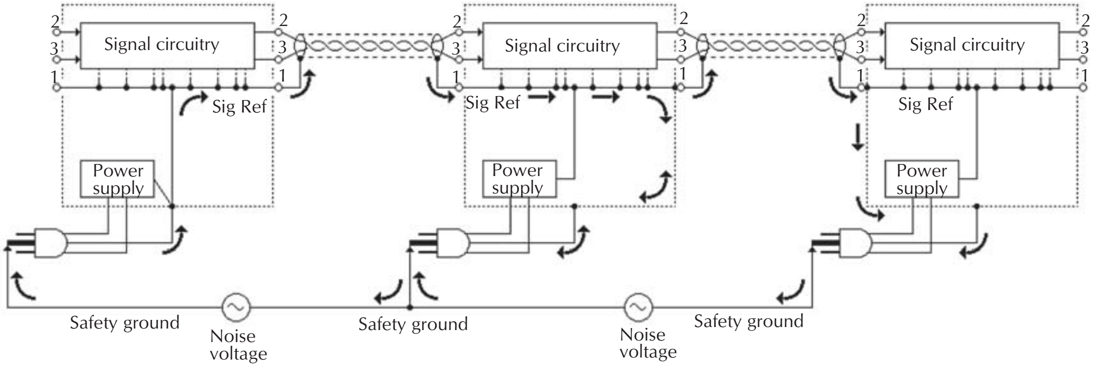
Figure 32-31. Old and new IEC tests for CMRR compared.

To the user, symptoms are indistinguishable from many other noise coupling problems such as poor CMRR. To quote Neil again,

> *Balancing is thus acquiring a tarnished reputation, which it does not deserve. This is indeed a curious situation. Balanced line-level interconnections are supposed to ensure noise-free system performance, but often they do not.*

In balanced interconnections, it occurs at line inputs and outputs where interconnecting cables routinely have their shields grounded at both ends. Of course, grounding at both ends is required for unbalanced interfaces.

Fig.32-32 illustrates several examples of common impedance coupling. When noise currents flow in signal reference wiring or circuit board traces, tiny voltage drops are created. These voltages can couple into the signal path, often into very high gain circuitry, producing hum or other noise at the output. In the first two devices, pin 1 current is allowed to flow in internal signal reference wiring. In the second and third devices, power line noise current (coupled through the parasitic capacitances in the power transformer) is also allowed to flow in signal reference wiring to reach the chassis/safety ground. This so-called *sensitive equipment* will produce additional noise independent of the pin 1 problem. For the second device, even disconnecting its safety ground (not recommended) won't stop current flow through it between input and output pin 1 shield connections.


2
3
1
Signal circuitry
2
3
1
Sig Ref
Power supply
2
3
1
Signal circuitry
2
3
1
Sig Ref
Power supply
2
3
1
Signal circuitry
2
3
1
Sig Ref
Power supply
Safety ground
Noise voltage
Safety ground
Noise voltage
Safety ground

A black dot at pin 1 indicates a direct chassis connection.

Figure 32-32. How poor routing of shield currents produces the pin 1 problem.

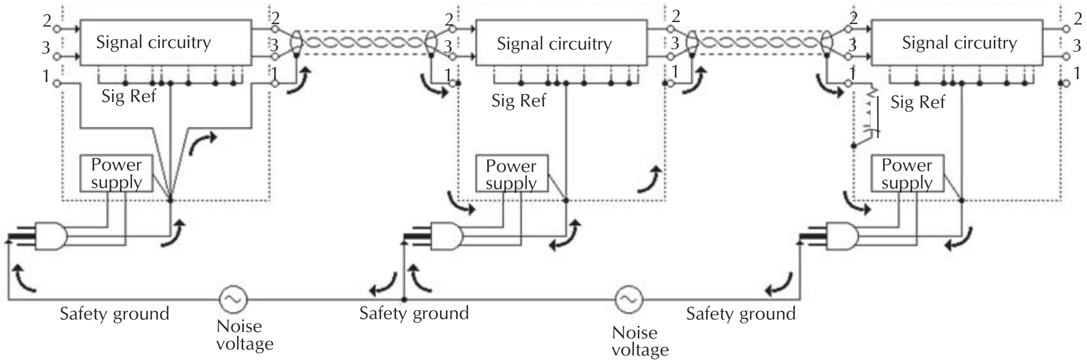


Figure 32-33. Equipment with proper internal grounding.

Fig. 32-33 shows three devices whose design does not allow shield current to flow in signal reference conductors. The first uses a star connection of input pin 1, output pin 1, power cord safety ground, and power supply common. This technique is the most effective prevention. Noise currents still flow, but not through internal signal reference conductors. Before there were printed circuit boards, a metal chassis served as a very low-impedance connection (effectively a ground plane) connecting all pins 1 to each other and to safety ground. Pin 1 problems were virtually unknown in those vintage designs. Modern printed circuit board–mounted connectors demand that proper attention be paid to the routes taken by ground noise currents. Of course, this same kind of problem can and does exist with RCA connectors in unbalanced consumer equipment, too.

Fortunately, tests to reveal such common impedance coupling problems are not complex. Comprehensive tests using lab equipment covering a wide frequency range have been described by Cal Perkins[31] and simple tests using an inexpensively built tester called the *hummer* have been described by John Windt.[32] Jensen Transformers, Inc. variant of the Hummer is shown in Fig. 32-34. It passes a rectified ac current of 60–80 mA through the potentially troublesome shield connections in the device under test to determine if they cause the coupling.

The glow of the automotive test lamp shows that a good connection has been made and that test current is indeed flowing. The procedure:

1. Disconnect all input and output cables, except the output to be monitored, as well as any chassis connections (due to rack mounting, for example) from the device under test.
2. Power up the device.

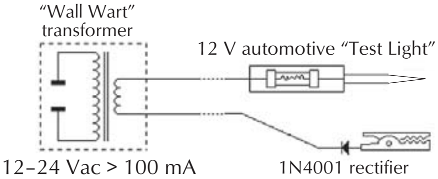


Figure 32-34. The Hummer II. Courtesy Jensen Transformers, Inc.

3. Meter and, if possible, listen to the device output. Hopefully, the output will simply be random noise. Try various settings of operator controls to familiarize yourself with the noise characteristics of the device under test without the hummer connected.
4. Connect the hummer clip lead to the device chassis and touch the probe tip to pin 1 of each input or output connector. If the device is properly designed, there will be no output hum or change in the noise floor.
5. Test other potentially troublesome paths, such as from an input pin 1 to an output pin 1 or from the safety ground pin of the power cord to the chassis (a three-to-two-prong ac adapter is handy to make this connection).

   Note: Pin 1 might not be connected directly to ground in some equipment—hopefully, this will be at inputs only! In this case, the hummer's lamp may not glow—this is OK.

32.5.3 Balanced Cable Issues

At audio frequencies, even up to about 1 MHz, cable shields should be grounded at one end only, where the signal is ground referenced. At higher frequencies, where typical system cables become a small fraction of a wavelength, it's necessary to ground it at more than one point to keep it at ground potential and guard against RF interference.[26,27] Based on my own work, there are two additional reasons that there should always be a shield ground at the driver end of the cable, whether the receiver end is grounded or not, see Figs. 32-35 and 32-36. The first reason involves the cable capacitances between each signal conductor and shield, which are mismatched by 4% in typical cable. If the shield is grounded at the receiver end, these capacitances and driver common-mode output impedances, often mismatched by 5% or more, form a pair of low-pass filters for common-mode noise. The mismatch in the filters converts a portion of common-mode noise to differential signal. If the shield is connected only at the driver, this mechanism does not exist. The second reason involves the same capacitances working in concert with signal asymmetry. If signals were perfectly symmetrical and capacitances perfectly matched, the capacitively coupled signal current in the shield would be zero through cancellation. Imperfect symmetry and/or capacitances will cause signal current in the shield. This current should be returned directly to the driver from which it came. If the shield is grounded at the receiver, all or part of this current will return via an undefined path that can induce crosstalk, distortion, or oscillation.[28]

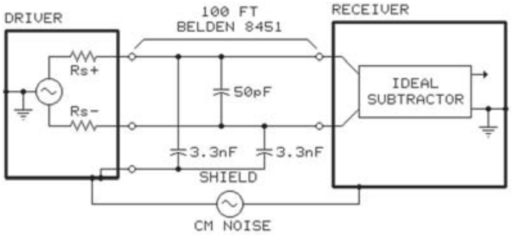


Figure 32-35. Shield grounded only at driver.

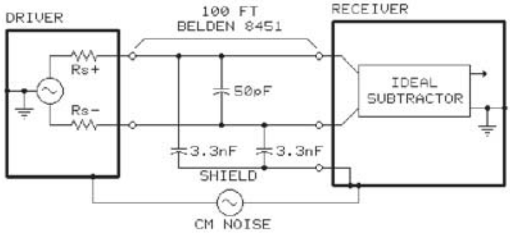


Figure 32-36. Shield grounded only at receiver.

With cables, too, there is a conflict between the star and mesh grounding methods as discussed in Section 32.4.3. But this low-frequency versus high-frequency conflict can be substantially resolved with a hybrid approach involving grounding the receive end of cables through an appropriate capacitance (shown in the third device of Fig. 32-33).[26,27] Capacitor values in the range of 10 nF to 100 nF are most appropriate for the purpose. Such capacitance has been integrated into the Neutrik EMC series connectors. The merits of this scheme were the subject of several years of debate in the Audio Engineering Society Standards Committee working group that developed AES48.

As discussed in Section 32.2.5, twisting essentially places each conductor at the same average distance from the source of a magnetic field and greatly reduces differential pickup. Star quad cable reduces pickup even further, typically by about 40 dB. But the downside is that its capacitance is approximately double that of standard shielded twisted pair.

SCIN, or shield-current induced noise, may be one consequence of connecting a shield at both ends. Think of a shielded twisted pair as a transformer with the shield acting as primary and each inner conductor acting as a secondary winding, as shown in the cable model of Fig. 32-37. Current flow in the shield produces a magnetic field which then induces a voltage in each of the inner conductors. If these voltages are identical, and the interface is properly impedanc balanced, only a common-mode voltage is produced that can be rejected by the line receiver. However, subtle variations in physical construction of the cable can produce unequal coupling in the two signal conductors. The difference voltage, since it appears as signal to the receiver, results in noise coupling. Test results on six commercial cable types appear in reference.[29] In general, braided shields perform better than foil shields and drain wires.

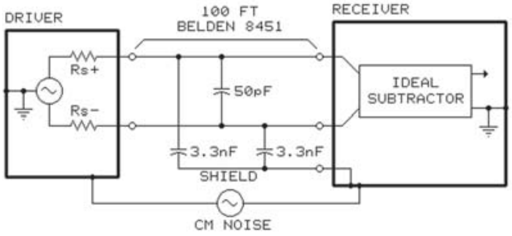


Figure 32-37. Shield of a shielded twisted pair cable is magnetically coupled to inner conductors.

And, to make matters even worse, grounding the shield of balanced interconnect cables at both ends also excites the pin 1 problem if it exists. Although it might appear that there's little to recommend grounding at both ends, it is a widely accepted practice. As you can see, noise rejection in a real-world balanced interface can be degraded by a number of subtle problems and imperfections. But, as discussed in Section 32.6.4, it is virtually always superior to an unbalanced interface!

32.5.4 Coupling in Unbalanced Cables

The overwhelming majority of consumer as well as high-end audiophile equipment still uses an audio interface system introduced over 60 years ago and intended to carry signals from chassis to chassis inside the earliest RCA TV receivers! The ubiquitous RCA cable and connector form an unbalanced interface that is extremely susceptible to common impedance noise coupling.

As shown in Fig. 32-38, noise current flow between the two device grounds or chassis is through the shield conductor of the cable. This causes a small but significant noise voltage to appear across the length of the cable. Because the interface is unbalanced, this noise voltage will be directly added to the signal at the receiver.[33] In this case, the impedance of the shield conductor is responsible for the *common impedance coupling*. This coupling causes hum, buzz, and other noises in audio systems. It's also responsible for slow-moving hum bars in video interfaces and glitches, lock-ups, or crashes in unbalanced—e.g., RS-232—data interfaces.

Consider a 25 ft interconnect cable with foil shield and a #26 AWG drain wire. From standard wire tables or actual measurement, its shield resistance is found to be 1.0 Ω. If the 60 Hz leakage current is 300 μA, the hum voltage will be 300 μV. Since the consumer audio reference level is about –10 dBV or 300 mV, the 60 Hz hum will be only $20\log(300\ \mu\text{V}/300\ \text{mV}) = -60\ \text{dB}$ relative to the signal. For most systems, this is a very poor signal-to-noise ratio! For equipment with two-prong plugs, the 60 Hz harmonics and other high-frequency power-line noise (refer to Fig. 32-20) will be capacitively coupled and result in a harmonic-rich buzz.

Because the output impedance of device A and the input impedance of device B are in series with the inner conductor of the cable, its impedance has an insignificant effect on the coupling and is not represented here. Common-impedance coupling can become extremely severe between two grounded devices, since the voltage drop in the safety ground wiring between the two devices is effectively parallel connected across the length of the cable shield. This generally results in a fundamental-rich hum that may actually be larger than the reference signal!

Coaxial cables, which include the vast majority of unbalanced audio cables, have an interesting and underappreciated quality regarding common-impedance coupling at high frequencies, Fig. 32-39. Any voltage appearing across the ends of the shield will divide itself between shield inductance L_s and resistance R_s according to frequency. At some frequency, the voltages across each will be equal (when reactance of L_s equals R_s). For typical cables, this frequency is in the 2 to 5 kHz range. At frequencies below this transition frequency, most of the ground noise will appear across R_s and be coupled into the audio signal as explained earlier. However, at frequencies above the transition frequency, most of the ground noise will appear across L_s. Since L_s is magnetically coupled to the inner conductor, a replica of the ground noise is induced over its length. This induced voltage is then subtracted from the signal on the inner conductor, reducing noise coupling into the signal. At frequencies ten times the transition frequency, there is virtually no noise coupling at all—common-impedance coupling has disappeared. Therefore, common-impedance coupling in coaxial cables ceases to be a noise issue at frequencies over about 50 kHz. Remember this as we discuss claims made for power line filters that typically remove noise only above about 50 kHz.

Unbalanced interface cables, regardless of construction, are also susceptible to magnetically induced noise caused by nearby low-frequency ac magnetic fields.

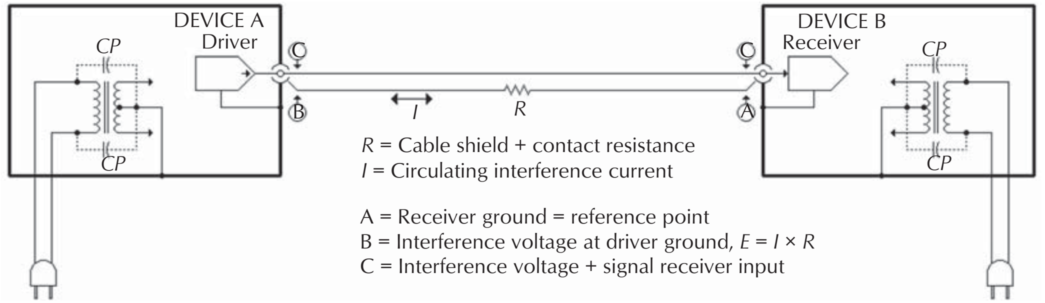

DEVICE A
Driver
CP
CP
C
B
I
R
C
A
DEVICE B
Receiver
CP
CP
R = Cable shield + contact resistance
I = Circulating interference current
A = Receiver ground = reference point
B = Interference voltage at driver ground, E = I × R
C = Interference voltage + signal receiver input

Figure 32-38. Common impedance coupling in an unbalanced audio, video, or data interface.

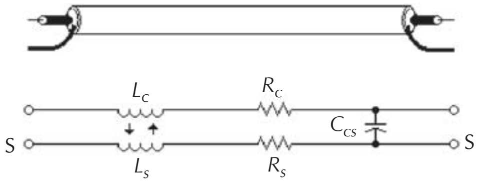


Figure 32-39. Magnetic coupling between shield and center conductor is 100%.

Unlike balanced interconnections, such noise pickup is not nullified by the receiver.

32.5.5 Bandwidth and RF Interference

RF interference isn't hard to find—it's actually very difficult to avoid, especially in urban areas. It can be radiated through the air and/or be conducted through any cables connected to equipment. Common sources of radiated RF include AM, shortwave, FM, and TV broadcasts; ham, CB, remote control, wireless phone, cellular phone, and a myriad of commercial two-way radio and radar transmitters; and medical and industrial RF devices. Devices that create electrical sparks, including welders, brush-type motors, relays, and switches can be potent wideband radiators. Less obvious sources include arcing or corona discharge from power line insulators (common in seashore areas or under humid conditions) or malfunctioning fluorescent, HID, or neon lights. Of course, lightning, the ultimate spark, is a well-known radiator of momentary interference to virtually anything electronic.

Interference can also be conducted via any wire coming into the building. Because power and telephone lines also behave as huge outdoor antennas, they are often teeming with AM radio signals and other interference. But the most troublesome sources are often inside the building and the energy delivered through ac power wiring. The offending source may be in the same room as your system or, worse yet, it may actually be a part of your system! The most common offenders are inexpensive light dimmers, fluorescent lights, CRT displays, digital signal processors, or any device using a switching power supply.

Although cable shielding is a first line of defense against RF interference, its effectiveness depends critically on the shield connection at each piece of equipment. Because substantial inductance is added to this connection by traditional XLR connectors and grounding pigtails, the shield becomes useless at high radio frequencies. Common-mode RF interference simply appears on all the input leads.[34] Because the wire limitations discussed in Section 32.2.4 apply to grounding systems, contrary to widespread belief, grounding is not an effective way to deal with RF interference. To quote Neil Muncy:

> *Costly technical grounding schemes involving various and often bizarre combinations of massive copper conductors, earth electrodes, and other arcane hardware are installed. When these schemes fail to provide expected results, their proponents are usually at a loss to explain why.*[35]

The wider you open the window, the more dirt flies in. One simple, but often overlooked, method of minimizing noise in a system is to limit the system bandwidth to that required by the signal.[36] In an ideal world, every signal-processing device in a system would contain a filter at each input and output connector to appropriately limit bandwidth and prevent out-of-band energy from ever reaching active circuitry. This RF energy becomes an audio noise problem because the RF is demodulated or detected by active circuitry in various ways, acting like a radio receiver that adds its output to the audio signal. Symptoms can range from actual reception of radio signals or a 59.94 Hz buzz from TV signals or various tones from cell phone signals to much subtler distortions, often described as a veiled or grainy audio quality.[37] The filters necessary to prevent these problems vary widely in effectiveness and, in some equipment, may not be present at all. Sadly, the performance of most commercial equipment will degrade when such interference is coupled to its input.[38]

32.6 Solving Real-World System Problems

How much noise and interference are acceptable depends on what the system is and how it will be used. Obviously, sound systems in a recording studio need to be much more immune to noise and interference than paging systems for construction sites.

32.6.1 Noise Perspective

The decibel is widely used to express audio-related measurements. For *power* ratios,

$$dB = 10\log\frac{P_1}{P_2} \qquad (32\text{-}9)$$

For *voltage* or *current* ratios, because power is proportional to the square of voltage or current:

$$dB = 20\log\frac{E_1}{E_2}$$

$$dB = 20\log\frac{I_1}{I_2} \qquad (32\text{-}10)$$

Most listeners describe 10 dB level decreases or increases as halving or doubling loudness, respectively, and 2 dB or 3 dB changes as just noticeable. Under laboratory conditions, well-trained listeners can usually identify level changes of 1 dB or less. The dynamic range of an electronic system is the ratio of its maximum undistorted signal output to its residual noise output or noise floor. Up to 120 dB of dynamic range may be required in high-end audiophile sound systems installed in typical homes.[39]

32.6.2 Troubleshooting

Under certain conditions, many systems will be acceptably noise-free in spite of poor grounding and interfacing techniques. People often get away with doing the wrong things! But, notwithstanding anecdotal evidence to the contrary, logic and physics will ultimately rule.

Troubleshooting noise problems can be a frustrating, time-consuming experience but the method described in Section 32.6.2.2 can relieve the pain. It requires no electronic instruments and is very simple to perform. Even the underlying theory is not difficult. The tests will reveal not only what the coupling mechanism is but also where it is.

32.6.2.1 Observations, Clues, and Diagrams

A significant part of troubleshooting involves how you think about the problem. First, don't assume anything! For example, don't fall into the trap of thinking, just because you've done something a particular way many times before, it simply can't be the problem. Remember, even things that can't go wrong, do! Resist the temptation to engage in guesswork or use a shotgun approach. If you change more than one thing at a time, you may never know what actually fixed the problem.

Second, ask questions and gather clues! If you have enough clues, many problems will reveal themselves before you start testing. Be sure to write everything down—imperfect recall can waste a lot of time! Troubleshooting guru Bob Pease[40] suggests these basic questions:

1. Did it ever work right?
2. What are the symptoms that tell you it's not working right?
3. When did it start working badly or stop working?
4. What other symptoms showed up just before, just after, or at the same time as the failure?

Operation of the equipment controls, and some elementary logic, can provide very valuable clues. For example, if a noise is unaffected by the setting of a gain control or selector, logic dictates that it must be entering the signal path after that control. If the noise can be eliminated by turning the gain down or selecting another input, it must be entering the signal path before that control.

Third, sketch a block diagram of the system. Fig. 32-40 is an example diagram of a simple home theater system. Show all interconnecting cables and indicate approximate length. Mark any balanced inputs or outputs. Generally, stereo pairs can be indicated with a single line. Note any device that is grounded via a three-prong ac plug. Note any other ground connections such as equipment racks, cable TV connections, etc.

32.6.2.2 The Ground Dummy Procedure

An easily constructed adapter or ground dummy is the key element in this procedure. By temporarily placing the dummy at strategic locations in the interfaces, precise information about the nature and location of the problem is revealed. The tests can specifically identify:

1. Common-impedance coupling in unbalanced cables.
2. Shield current-induced coupling in balanced cables.
3. Magnetic or electrostatic pickup of nearby magnetic or electrostatic fields.
4. Common-impedance coupling (the pin 1 problem) inside defective devices.
5. Inadequate CMRR of the balanced input.

The ground dummy can be made from standard connector wired as shown in Figs. 32-41 and 32-42. Since a dummy does not pass signal, mark it clearly to help prevent it, being accidentally left in a system.

Each signal interface is tested in four steps. As a general rule, always start at the inputs to the power amplifiers and work backward toward the signal sources. *Be very careful when performing the tests not to damage loudspeakers or ears!* The surest way to avoid possible damage is to turn off the power amplifier(s) before reconfiguring cables for each test step.

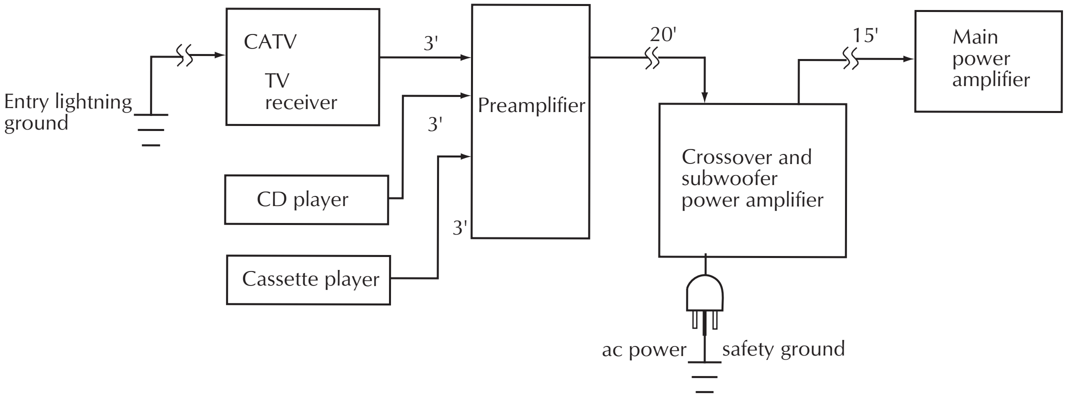
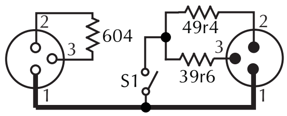

Entry lightning ground
CATV
TV receiver
3'
CD player
3'
Cassette player
3'
Preamplifier
20'
15'
Main power amplifier
Crossover and subwoofer power amplifier
ac power
safety ground

Figure 32-40. Block diagram of example system.


2
3
604
49r4
2
3
39r6
S1
1
1

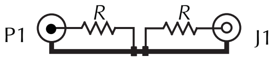
For balanced audio XLR
P1/J1 = Switchcraft S3FM adapter
with QG3F and QG3M insert
All resistors 1%, ¼ W metal film

Close S1 for CMRR test only

For balanaced audio 3C phone
Use Switchcraft 383A and 387A
adapters with XLR version

Figure 32-41. Balanced ground dummy.

Unbalanced interfaces



For audio RCA
P1 = Switchcraft 3502 plug
J1 = Switchcraft 3503 jack
R = 1 kΩ, 5%, ¼ W resistor

For audio 2C phone
Use Switchcraft 336A and 345A
adapters with RCA version

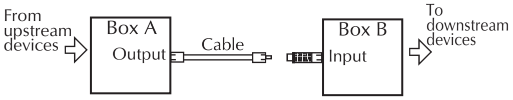
Figure 32-42. Unbalanced ground dummy.

32.6.2.2.1 For Unbalanced Interfaces

STEP 1: Unplug the cable from the input of Box B and plug in only the dummy as shown below.

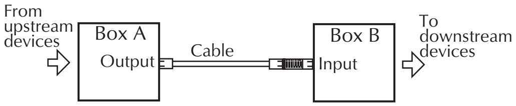
- ♦ Output quiet?

  No—The problem is either in Box B or farther downstream.


From upstream devices
Box A
Output
Cable
Box B
Input
To downstream devices

  Yes—Go to next step.

STEP 2: Leaving the dummy in place at the input of Box B, plug the cable into the dummy as shown below.

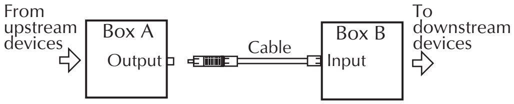

From upstream devices
Box A
Output
Cable
Box B
Input
To downstream devices

- ♦ Output quiet?

  No—Box B has a pin 1 problem (see Section 4.3 to confirm this).

  Yes—Go to next step.

STEP 3: Remove the dummy and plug the cable directly into the input of Box B. Unplug the other end of the cable from the output of Box A and plug it into the dummy as shown below. Do not plug the dummy into Box A or let it touch anything conductive.

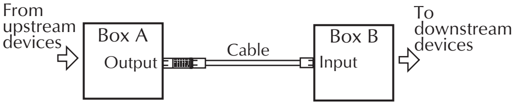

From upstream devices
Box A
Output
Cable
Box B
Input
To downstream devices

- ♦ Output quiet?

  No–Noise is being induced in the cable itself. Reroute the cable to avoid interfering fields (see Section 32.4.2 or 32.4.4).

  Yes—Go to next step.

STEP 4: Leaving the dummy in place on the cable, plug the dummy into the output of Box A as shown below.


From
upstream
devices
Box A
Output
Cable
Box B
Input
To
downstream
devices

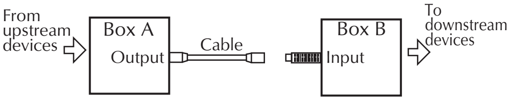
- ♦ Output quiet?

No—The problem is common-impedance coupling (see Section 32.4.4). Install a ground isolator at the input of Box B.

Ye—The noise is coming from (or through) the output of Box A. Perform the same test sequence on the cable(s) connecting Box A to upstream devices.

32.6.2.2.2 For Balanced Interfaces

STEP 1: Unplug the cable from the input of Box B and plug in only the dummy (switch open or NORM) as shown below.

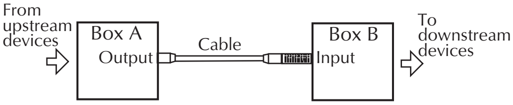

From
upstream
devices
Box A
Output
Cable
Box B
Input
To
downstream
devices

- ♦ Output quiet?

No—The problem is either in Box B or farther downstream.

Yes—Go to next step.

STEP 2: Leaving the dummy in place at the input of Box B, plug the cable into the dummy (switch open or NORM) as shown below.

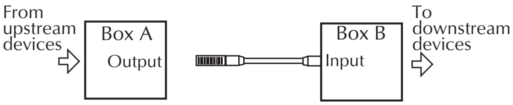

From
upstream
devices
Box A
Output
Cable
Box B
Input
To
downstream
devices

- ♦ Output quiet?

N—Box B has a Pin 1 problem (see hummer test, Section 32.4.2, to confirm this).

Yes—Go to next step.

STEP 3: Remove the dummy and plug the cable directly into the input of Box B. Unplug the other end of the cable from the output of Box A and plug it into the dummy (switch open or NORM) as shown below. *Do not plug the dummy into Box A or let it touch anything conductive.*

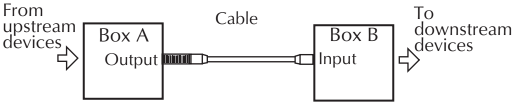

From
upstream
devices
Box A
Output
Box B
Input
To
downstream
devices

- ♦ Output quiet?

No—Noise is being induced in the cable itself by an electric or magnetic field. Check the cable for an open shield connection, reroute the cable to avoid the interfering field, or replace the cable with a starquad type (see Sections 32.2.5 and 32.4.3).

Yes—Go to next step.

STEP 4: Leaving the dummy in place on the cable, plug the dummy (switch open or NORM) into the output of Box A as shown below.


From
upstream
devices
Box A
Output
Cable
Box B
Input
To
downstream
devices

- ♦ Output quiet?

No—The problem is shield-current-induced noise (see Section 32.4.3). Replace the cable with a different type (without a drain wire) or take steps to reduce current in the shield.

Yes—Go to next step.

STEP 5: Leave the dummy and cable as for step 4, but move the dummy switch to the CMRR (closed) position.

- ♦ Output quiet?

No—The problem is likely inadequate common-mode rejection of the input stage of Box B. This test is based on the IEC common-mode rejection test but uses the actual common-mode voltage present in the system. The nominal 10 Ω imbalance may not simulate the actual imbalance at the output of Box A, but the test will reveal input stages whose CMRR is sensitive to source imbalances. Most often, adding a transformer-based ground isolator at the input of Box B will cure the problem.

Yes—The noise must be coming from (or through) the output of Box A. Perform the same test sequence on the cable(s) connecting Box A to upstream devices.

32.6.3 Solving Interface Problems

32.6.3.1 Ground Isolators

A device called a *ground isolator* solves the inherent common-impedance coupling problem in unbalanced interfaces. Broadly defined, a ground isolator is a differential responding device with high common-mode rejection. It is not a filter that can selectively remove hum, buzz, or other noises when simply placed anywhere in the signal path. To do its job, it must be installed where the noise coupling would otherwise occur.

A transformer is a passive device that fits the definition of a ground isolator. Transformers transfer a voltage from one circuit to another without any electrical connections between the two circuits. It converts an ac signal voltage on its primary winding into a fluctuating magnetic field that is then converted back to an ac signal voltage on its secondary winding (discussed in detail in Chapter 11).

As shown in Fig. 32-43, when a transformer is inserted into an unbalanced signal path, the connection between device grounds via the cable shield is broken. This stops the noise current flow in the shield conductor that causes the noise coupling, as discussed in Section 32.5.4. As discussed in Chapter 11, the highest noise rejection is achieved with input-type transformers containing Faraday shields. A transformer-based isolator for consumer audio signals using such transformers, the ISO-MAX® model CI-2RR, is shown in Fig. 32-44. To avoid bandwidth loss, such isolators must be located at the receive end of interconnections, using minimum-length cables between isolator outputs and equipment inputs. Conversely, isolators using output-type transformers, such as the ISO-MAX® model CO-2RR and most other commercial isolators, may be freely located but will achieve significantly less noise rejection.

Ground isolators can also solve most of the problems associated with balanced interfaces. The ISO-MAX® Pro model PI-2XX shown in Fig. 32-45 often improves CMRR by 40 dB to 60 dB and provides excellent CMRR even if the signal source is unbalanced. Because it also features DIP switches to reconfigure cable shield ground connections, it can also solve pin 1 problems. Because it uses input-type transformers, it attenuates RF interference such as AM radio by over 20 dB. Again, to avoid bandwidth loss, it must be located at the receive end of long cable runs, using minimum-length cables between isolator outputs and equipment inputs. Other models are available for microphone signals and other applications. The vast majority of commercial hum eliminators and a few special-purpose ISO-MAX® models use output-type transformers, which may be freely located but offer significantly less CMRR improvement and have essentially no RF attenuation.

Several manufacturers make active (i.e., powered) ground isolators using some form of the simple differential amplifier shown in Fig. 32-31. Unfortunately, these circuits are exquisitely sensitive to the impedance of the driving source. Fig. 32-46 compares the measured 60 Hz (hum) rejection of a typical active isolator to a transformer-based isolator. Over the typical range of consumer output impedances, 100 Ω to 1 kΩ, the transformer has about 80 dB more rejection!

Passive isolators based on input-type transformers have other advantages, too. They require no power, they inherently suppress RF interference, and they're immune to most overvoltages that can be sudden death to active circuitry.

32.6.3.2 Multiple Grounding

When a system contains two or more grounded devices, such as the TV receiver and the subwoofer power amplifier in our example home theater system, a wired ground loop is formed as shown in Fig. 32-47.

As discussed in Sections 32.5.3 and 32.5.4, noise current flowing in the shaded path can couple noise into the signal as it flows in unbalanced cables or through the equipment's internal the ground path. This system would likely exhibit a loud hum regardless of the input selected or the setting of the volume control because of noise current flow in the 20 ft cable. You might be tempted to break this ground loop by lifting the safety ground at the subwoofer. Reread Section 32.4.2 and don't do it!

One safe solution is to break the ground loop by installing a ground isolator in the audio path from preamp to subwoofer as shown in Fig. 32-48. This isolator could also be installed in the path from TV receiver to preamp, but it is generally best to isolate the longest lines since they are more prone to coupling than shorter ones.

Another safe solution is to break the ground loop by installing a ground isolator in the CATV signal path at the TV receiver as shown in Fig. 32-49. These RF isolators generally should be installed where the cable connects to the local system, usually at a VCR or TV input. If an RF isolator is used at the input to a splitter, ground loops may still exist between systems served by the splitter outputs since the splitter provides no ground isolation. Although it can be used with a conventional TV or FM antenna, never install an RF isolator between

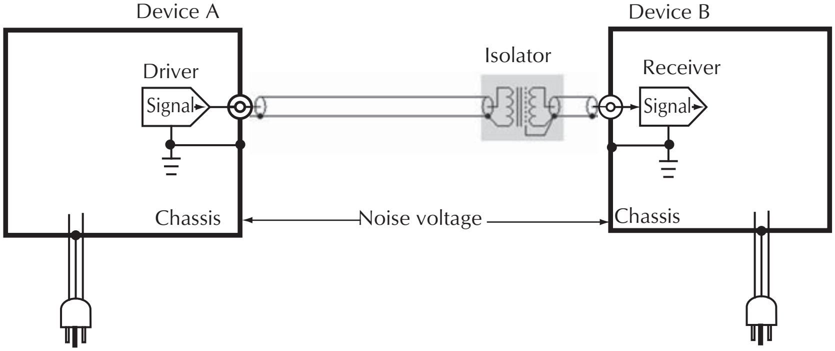


Figure 32-43. Ground isolator stops noise current in shield of unbalanced cable.

ISO-MAX
Made in U.S.A. by JENSEN TRANSFORMERS, INC.
L
R
IN
L
R
OUT

Figure 32-44. Stereo unbalanced audio isolator. Courtesy of Jensen Transformers, Inc.

ISO-MAX
IN
CH 1
OUT
IN
CH 2
OUT

Figure 32-45. Stereo balanced audio isolator. Courtesy of Jensen Transformers, Inc.

the CATV drop or antenna and its lightning ground connection (see Section 32.4.1). Isolators will *not* pass dc operating power to the dish in DBS TV systems.

Since most unbalanced interfaces are made to consumer devices that have two-prong ac plugs, isolating the signal interfaces may leave one or more pieces of equipment with no ground reference whatsoever. This could allow the voltage between an isolator's input and output to reach 50 Vac or more. While this isn't dangerous (leakage current is limited in UL-listed devices), it would require unrealistically high (CMRR over 140 dB) performance by the isolator to reject it! The problem is solved by grounding any floating gear as shown in Fig. 32-50. This is best done by replacing the two-prong ac plug with a three-prong type and adding a wire (green preferred) wire connected between the safety ground pin of the new ac plug and a chassis ground point.

A screw may be convenient as the chassis ground point. Use an ohmmeter to check for continuity between the screw and the outer contact of an RCA connector, which itself can be used if no other point is available. Although, in the example above, an added ground at either the preamp or the power amp would suffice, grounding the device with the highest leakage current—usually those devices with the highest ac power consumption rating—will generally result in the lowest noise floor.

32.6.3.3 Unbalanced to Balanced and Vice Versa

The reader is referred to Chapter 11, Section 11.2.2, Audio Transformers, for a more detailed discussion of these applications.

Beware of RCA to XLR adapters! Fig. 32-51 shows how using this adapter to connect an unbalanced output to a balanced input reduces the interface to an unbalanced one having absolutely *no* ground noise rejection! The potential noise reduction benefit of the balanced input is completely lost.

Proper wiring for this interface, shown in Fig. 32-52, results in at least 20 dB of noise rejection even if the balanced input is one of typically mediocre performance. The key difference is that, by using shielded twisted pair cable, the ground noise current flows in a separate conductor that is *not* part of the signal path.

Driving an unbalanced input from a balanced output is not quite as straightforward. Balanced equipment outputs use a wide variety of circuits. Some, such as the one in Fig. 32-53, might be damaged when one output is grounded. Others, including most popular servobalanced output stages, can become unstable unless the output is

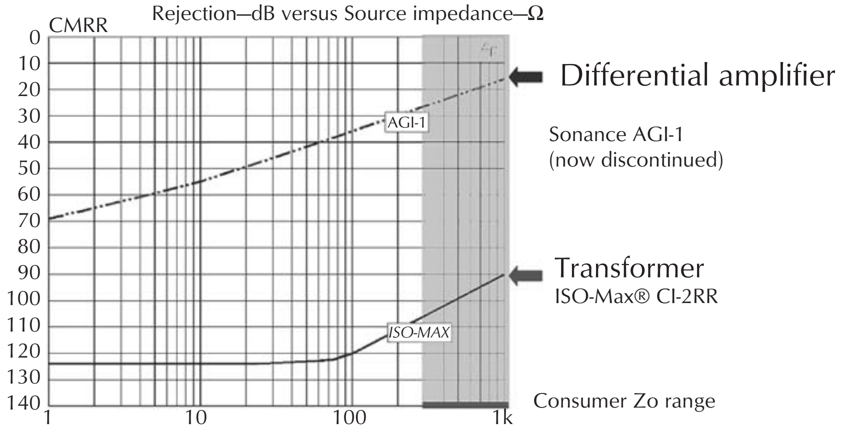


Figure 32-46. Measured hum rejection versus source impedance, active differential amplifier versus input transformer isolator.

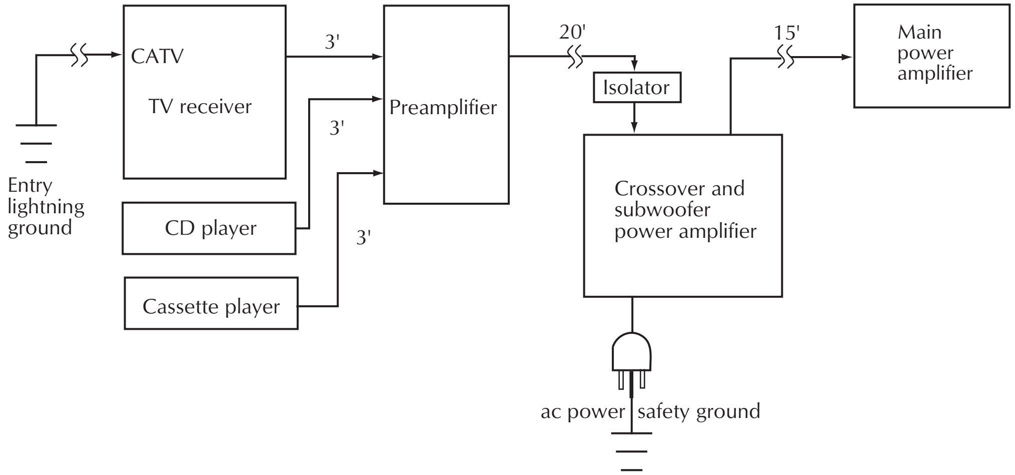

CATV
TV receiver
Entry
lightning
ground
CD player
Cassette player
3'
3'
3'
Preamplifier
20'
Isolator
Crossover and
subwoofer
power amplifier
15'
Main
power
amplifier
ac power
safety ground

Figure 32-47. Using an audio ground isolator to break the loop.

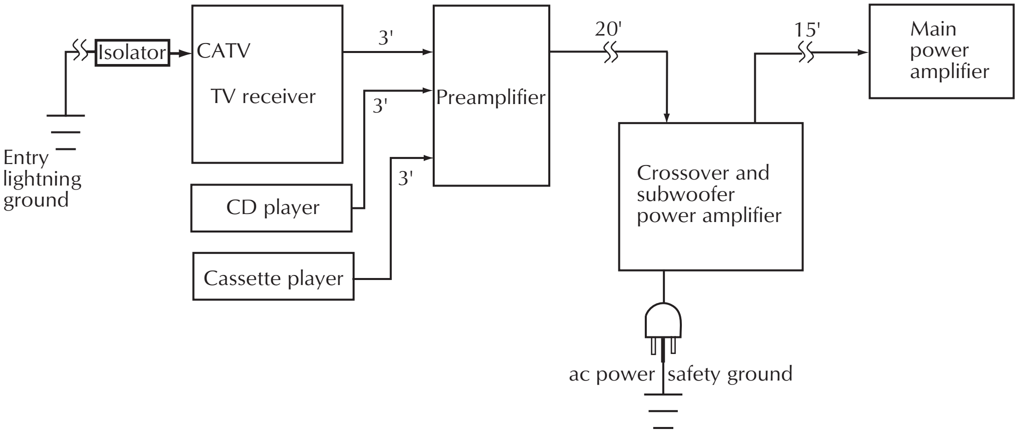

Isolator
CATV
TV receiver
Entry
lightning
ground
CD player
Cassette player
3'
3'
3'
Preamplifier
20'
Crossover and
subwoofer
power amplifier
15'
Main
power
amplifier
ac power
safety ground

Figure 32-48. Using a CATV ground isolator to break the loop.

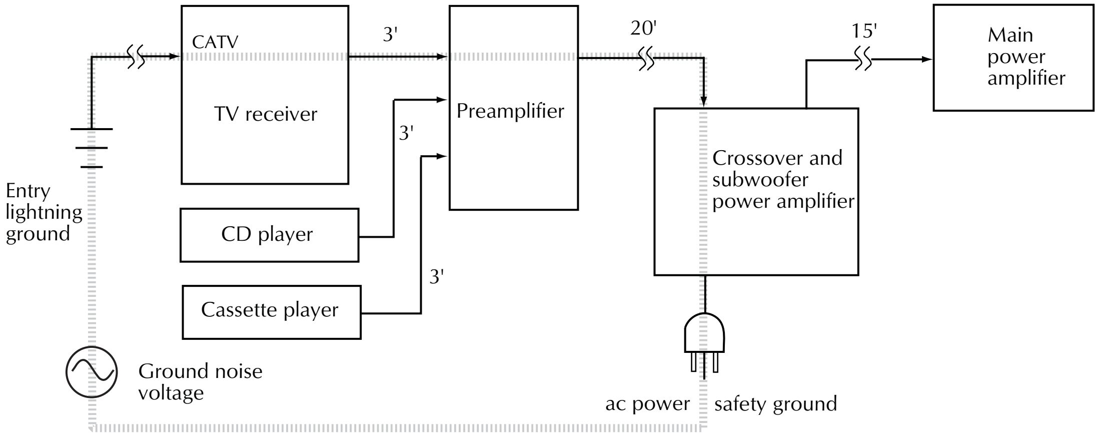

CATV
TV receiver
3'
Preamplifier
3'
3'
20'
15'
Main power amplifier
Crossover and subwoofer power amplifier
Entry lightning ground
CD player
Cassette player
Ground noise voltage
ac power
safety ground

Figure 32-49. Loop created by two ground connections.

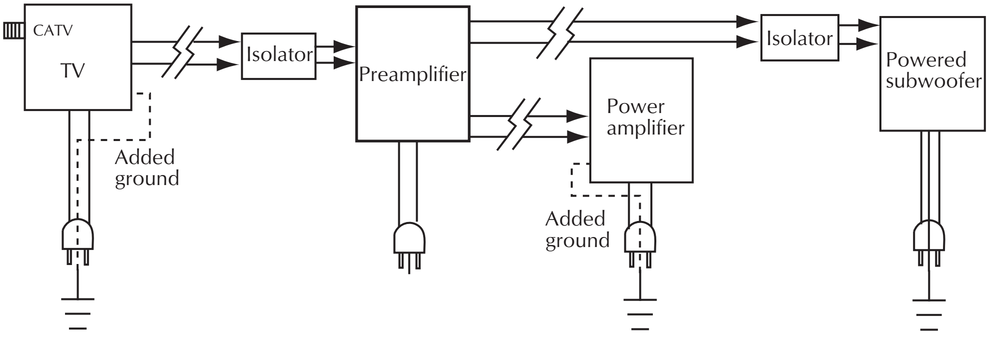

CATV
TV
Isolator
Preamplifier
Isolator
Powered subwoofer
Power amplifier
Added ground
Added ground

Figure 32-50. Grounding floating equipment when isolators are installed. From Jensen AN004.

directly grounded at the driver end, which reduces the interface to an unbalanced one with no noise rejection.[41] Unless a balanced output already utilizes a built-in transformer, using an external ground isolator such as the one shown in Fig. 32-53 is the only method that will simultaneously avoid weird or damaging behavior and minimize ground noise when used with virtually any output stage. This approach is used in the ISO-MAX® Pro model PC-2XR pro-to-consumer interface.

32.6.3.4 RF Interference

As mentioned earlier, immunity to RF interference or RFI is part of good equipment design. Testing for RFI susceptibility is now mandated in Europe. Unfortunately, much of the equipment available today may still have very poor immunity. Under unfavorable conditions, external measures may be needed to achieve adequate immunity.[42]

For RF interference over about 20 MHz, ferrite clamshell cores shown in Fig. 32-54, which are easily installed over the outside of a cable, can be very effective. Some typical products are Fair-Rite #0431164281 and Steward #28A0640-0A.[43,44] In most cases, they work best when placed on the cable at or near the receive end. Often they are more effective if the cable is looped through the core several times.

If this is inadequate, or the frequency is lower (such as AM radio) you may have to add a low-pass—i.e., high-frequency reject—RFI filter on the signal line. Fig. 32-55 shows sample 50 kHz cutoff, 12 dB per octave low-pass RFI filters for unbalanced or balanced audio applications. For best performance and audio quality, use NP0 (also called C0G)-type ceramic capacitors keeping leads as short as possible, under ¼ inch preferred. For stubborn AM radio interference, it may help to increase the value of C up to about 1000 pF maximum. The 680 µH inductors are small ferrite core types such as J.W. Miller 78F681J or Mouser 434-22-681. If the only interference is above about 50 MHz, a ferrite bead may be used for L. For the balanced filter, inductors and capacitors should be ±5%

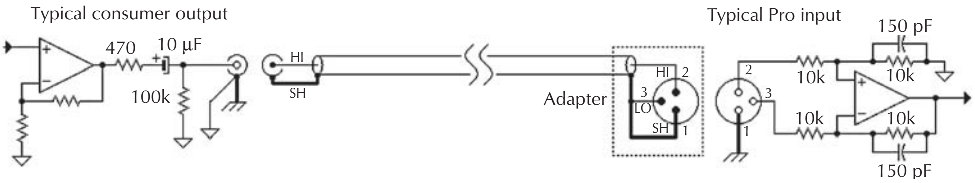

Typical consumer output
470
10 µF
100k
HI
SH
Adapter
HI
2
3
LO
SH
1
Typical Pro input
2
3
1
10k
10k
10k
10k
150 pF
150 pF

Figure 32-51. Incorrect connection of an unbalanced output to a balanced input.

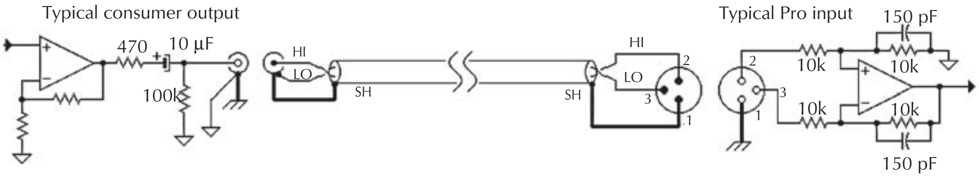

Typical consumer output
470
10 µF
100k
HI
LO
SH
SH
HI
LO
2
3
1
Typical Pro input
2
3
1
10k
10k
10k
10k
150 pF
150 pF

Figure 32-52. Correct connection of an unbalanced output to a balanced input.

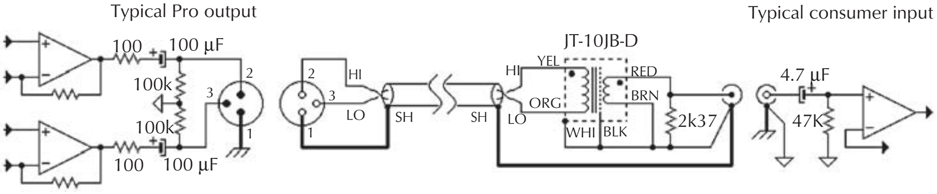

Typical Pro output
100
100 µF
100k
100k
100
100 µF
2
3
1
2
3
1
HI
LO
SH
JT-10JB-D
YEL
HI
ORG
SH
LO
WHI
BLK
RED
BRN
2k37
Typical consumer input
4.7 µF
47K

Figure 32-53. Bullet-proof connection of balanced output to unbalanced input. Use the shortest possible cable on the RCA side to avoid common-impedance coupling.

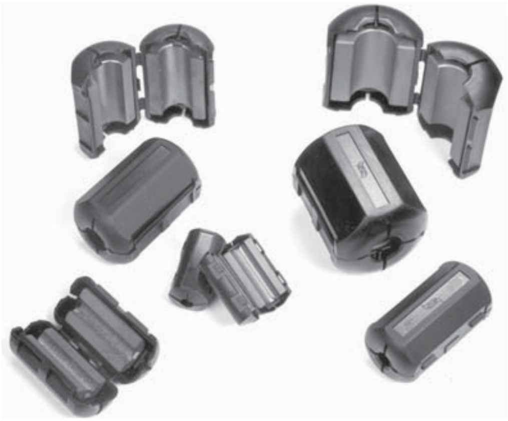


Figure 32-54. Ferrite cores.

tolerance parts or better to maintain impedance balance. The balanced filter can be used for low-level microphone lines, but miniature toroidal inductors are recommended to minimize potential hum pickup from stray magnetic fields. These filters, too, are generally most effective at the receive end of the cable.

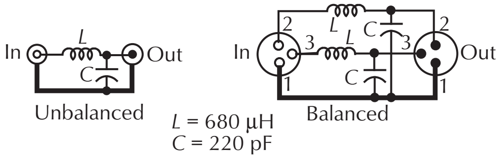

In
L
C
Out
Unbalanced
In
2
3
1
L
L
C
C
3
2
1
Out
Balanced
L = 680 µH
C = 220 pF

Figure 32-55. RF interference filters for audio lines.

When possible, the best way to deal with RF interference is to control it at its source. Fig. 32-56 is a schematic of a simple interference filter for solid-state 120 Vac light dimmers rated up to 600 W. It begins attenuating at about 50 kHz and is quite effective at suppressing AM radio interference. It must be installed within a few inches of the dimmer and, unfortunately, the components are large enough that it usually requires an extra-deep or a double knock-out box to accommodate both dimmer and filter. Parts cost is under $10.

A speaker cable can also become an antenna. In a strong RF field, enough voltage can be delivered to the semiconductors in the power amplifier that they become a detector and interference can be heard in the speaker even though the amplifier may be unpowered! More commonly, this problem occurs when the amplifier is powered and RF enters its feedback loop. In either case, the solution depends on the frequency of the interference. Ferrite cores on the cable near the amplifier may

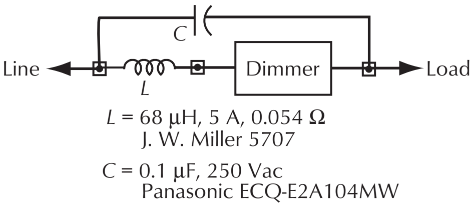


Figure 32-56. RF interference filter for solid-state light dimmer.

help. In stubborn cases, a 0.1 μF or a 0.22 μF capacitor directly across the output terminals of the amplifier may also be required.

32.6.3.5 Signal Quality

Audio transformer design involves a set of complex tradeoffs. The vast majority of available audio transformers, even when used as directed, fall short of professional performance levels. As Cal Perkins once wrote, "With transformers, you get what you pay for. Cheap transformers create a host of interface problems, most of which are clearly audible."[45]

The frequency response of a high-quality audio transformer is typically ruler flat, ±0.1 dB from 20 Hz to 20 kHz and –3 dB at 0.5 Hz and 100 kHz. The extended low-frequency response is necessary to achieve low phase distortion.[46] The high-frequency response is tailored to fall gradually, following a Bessel function. This, by definition, eliminates overshoot on square waves and high-frequency response peaking. Dramatic improvements in sonic clarity due to the Bessel filter action are often reported by Jensen customers who add transformers at power amplifier inputs. On the other hand, cheap transformers often have huge ultrasonic peaks in their response that are known to excite particularly ugly intermodulation distortions in even the finest downstream power amplifiers.[47]

Accurate *time domain* performance, sometimes called *transient response*, requires low phase distortion to preserve musical timbre and maintain accurate stereo imaging. Phase distortion not only alters sonic quality, it can also have serious system head room effects. Even though it may have a flat frequency response, a device having high phase distortion can increase peak signal amplitudes up to 15 dB. Phase distortion should never be confused with phase shift. Linear phase shift with frequency is simply a benign time delay: only deviations from linear phase or DLP create true phase distortion.[48] This DLP in a high-quality audio transformer is typically under 2° across the entire audio spectrum.

Harmonic and intermodulation distortion in audio transformers is unusually benign in character and cannot fairly be compared to electronic distortion. By their nature, transformers produce the most distortion when driven at high levels at very low frequencies, where the major distortion product is third harmonic. Transformer distortion mechanisms are frequency selective in a way that amplifiers, for example, are not. Electronic nonlinearities tend to produce harmonic distortions that are constant with frequency while high-quality transformer harmonic distortions drop to well under 0.001% at frequencies over a few hundred Hz. Transformers also tend to have remarkably low intermodulation distortion or IMD, to which the ear is particularly sensitive. Compared to an amplifier of comparable low-frequency harmonic distortion, a transformer typically has only a tenth the IMD. While cheap audio transformers use steel cores producing 1% low-frequency harmonic distortion at any signal level, high-quality transformers use cores of special nickel-iron-molybdenum alloys for vanishingly low distortion.

Of course, noise rejection or CMRR is often the most important property of a ground isolator. As discussed in Section 32.6.3.1 and Chapter 11, a transformer requires an internal Faraday shield (not a magnetic or case shield) to maximize CMRR. Most commercial isolators or hum eliminators consist of tiny imported telephone-grade transformers that do not contain such a shield. Beware of products with vague or nonexistent specs! For example, distortion described as under 0.1% is meaningless because frequency, signal level, and source impedance are not specified. The most common problems with inexpensive isolators are marginal noise reduction, loss of deep bass, bass distortion, and poor transient response. Of course, ad copy and specifications of these transformers will put on their best face, withholding the ugly truth! However, isolators using well-designed and properly applied audio transformers qualify as true high-fidelity devices. They are passive, stable, reliable, and require neither trimming, tweaking, nor excuses.

32.6.3.6 Tips for Balanced Interfaces

Be sure all balanced line pairs are twisted. Twisting is what makes a balanced line immune to interference from magnetic fields. This is especially important in low-level microphone cabling. Wiring at terminal or punch-down blocks and XLR connectors is vulnerable because the twisting is opened up, effectively creating a magnetic pickup loop. In very hostile environments, consider starquad cable because it has less susceptibility to magnetic fields. Magnetic coupling is also reduced by separation distance, cables crossing at right angles rather than running parallel, and shielding with magnetic material such as steel EMT conduit.

Pay attention to cable shield grounding. As discussed in Section 32.5.3, the shield must be grounded at the driven end, it may be grounded at both ends, but never grounded only at the receive end. As a standard practice, grounding at both ends is recommended for two reasons:

1. If the device input has marginal RF suppression, grounding the shield at the input will usually reduce problems,
2. It doesn't require the use of a specially wired cable that might find its way into another system and cause unexpected problems. If special cables are made—to deal with a pin 1 problem, for example—be sure they are clearly marked.

Don't terminate to reduce noise. Nearly every practical audio system should use unterminated audio circuits. This is standard professional audio practice worldwide. While a 600 Ω termination resistor at an input may reduce noise by up to 6 dB or more, depending on the driver output impedance, it will also reduce the signal by the same amount, so nothing is gained. If noise is caused by RF interference, installation of a suitably small capacitor at the input may be much more appropriate.

Use ground isolators to improve noise rejection. As discussed in Section 32.4.1, common balanced input circuits have generally unpredictable noise rejection in real-world systems. Actual in-system CMRR can be as little as 30 dB when using balanced sources and as little as 10 dB when using unbalanced sources. A quality transformer-based ground isolator can increase the CMRR of even the most mediocre balanced input to over 100 dB.

Beware of the pin 1 problem. As much as 50% of commercial equipment, some from respected manufacturers, has this designed-in problem. If disconnecting the shield at an input or output reduces a hum problem, the device at one or the other end of that cable may be the culprit. See Section 32.5.3 for test methods. Loose connector-mounting hardware is a major cause of pin 1 problems. Never overlook the obvious!

32.6.3.7 Tips for Unbalanced Interfaces

Keep cables as short as possible. Longer cables increase the coupling impedance. Serious noise coupling is nearly certain with 50 ft or 100 ft cables. Even much shorter cables can produce severe problems if there are multiple grounds. And never coil excess cable length.

Use cables with heavy gauge shields. Cables with shields of foil and thin drain wires increase the common-impedance coupling. Use cables with heavy braided copper shields, especially for long cables. See Section 32.7.4 for a recommended high-performance cable. The only property of cable that has any significant effect on audio-frequency noise coupling is shield resistance, which can be measured with an ordinary ohmmeter.

Bundle signal cables. All signal cables between any two boxes should be bundled. For example, if the *L* and *R* cables of a stereo pair are separated, nearby ac magnetic fields will induce a current in the loop formed by the two shields, causing hum in both signals. Likewise, all ac power cords should be bundled. This will tend to average and cancel the magnetic and electrostatic fields they radiate. In general, keeping signal bundles and power bundles separated will reduce coupling.

Maintain good connections. Connectors left undisturbed for long periods can oxidize and develop high contact resistance. Hum or other interference that changes when the connector is wiggled indicates a poor contact. Use a good commercial contact fluid and/or gold-plated connectors to help prevent such problems.

Don't add unnecessary grounds! Additional grounding almost always *increases* circulating noise currents rather than reducing them. As emphasized earlier, *never* disconnect or defeat the purpose of safety or lightning ground connections to solve a noise problem—the practice is both illegal and very dangerous!

Use ground isolators at problem interfaces. Transformer-based isolators magnetically couple the signal while completely breaking the noise current path through the cable and connectors. This eliminates common-impedance coupling and can improve immunity to RF interference as well.

Predict and solve problems before an installation. For systems that consist mostly of devices with two-prong power cords, some very simple multimeter measurements on each system device and cable makes it possible to actually predict hum levels and identify the problem interfaces before a system is installed.[49]

32.7 Alternative Treatments and Pseudoscience

The audio industry, especially the high-end segment, abounds with misinformation and myth. Science, evidence, and common sense are often discarded in favor of mysticism, marketing hype, and huge profits. Just remember that *the laws of physics have not changed!* See Fig. 32-57.

EINSTEIN
PHYSICS
POLICE

Figure 32-57. Officer Einsten of the Physics Police. Courtesy Coil-Craft.

32.7.1 Grounding from Technical to Bizarre

In most commercial buildings, the ac outlets on any branch circuit are saddle grounded or SG-types mounted in metallic J-boxes. Since SG outlets connect their safety ground terminals to the J-box, the safety ground network may now be in electrical contact with plumbing, air ducts, or structural building steel. This allows coupling of noisy currents from other loads—which might include air conditioning, elevators, and other devices—into the ground used by the sound system. In a scheme called *technical* or *isolated* grounding, safety grounding is not provided by the J-box and conduit but by a separate insulated green wire that must be routed back to the electrical panel alongside the white and black circuit conductors to keep inductance low. The technique uses special insulated ground or IG outlet—marked with a green triangle and sometimes orange in color—which intentionally insulates the green safety ground terminal from the outlet mounting yoke or saddle. The intent of the scheme is to isolate safety ground from conduit. Noise reduction is sometimes further improved by wiring each outlet as a home run back to the electrical panel or subpanel, making each outlet essentially a separate branch circuit.[50] This technique is covered by NEC Article 250-74 and its exceptions. In many cases, simply adding a new branch circuit can be just as effective yet far less expensive than implementing a technical ground system.

Many people, believing that the earth simply absorbs noise, have a strong urge to install multiple earth ground rods to fix noise. This is desperation-mode thinking. Code allows extra ground rods, but only if they are bonded to an existing properly implemented safety ground system. Code does not allow them to be used as a substitute soil resistance is simply too high and unstable to be relied on to divert fault currents.[51]

Equipment grounding via the standard power cord safety ground is logical, easy to implement, and safe. It's highly recommended for all systems and is the only practical method for portable or often reconfigured systems.

32.7.2 Power-Line Isolation, Filtering, and Balancing

Most sound systems use utility ac power. If it is disconnected, of course, all hum and noise disappears. This often leads to the odd conclusion that the noise is brought in with the power and that the utility company or the building wiring is to blame.[52] Devices claiming to cleanse and purify ac power have great intuitive appeal and are often applied without hesitation or much thought. A far more effective approach is to locate, and safely eliminate, ground loops that cause coupling of noise into the signal. This solves the real problem. In reality, when system designs are correct, special power treatment is rarely necessary. *Treating the power line to rid it of noise is analogous to repaving all the highways to fix the rough ride of a car. It's much more sensible to correct the cause of the coupling by replacing the shock absorbers!*

First, when any cord-connected line filter, conditioner, or isolation transformer is used, Code requires that the device as well as its load still be connected to safety ground as shown in Fig. 32-58. Cord-connected isolation transformers cannot be treated as separately derived sources unless they are permanently wired into the power distribution system per Code requirements. Sometimes makers of isolation transformers have been known to recommend grounding the shield and output to a separate ground rod. Not only does this violate Code, but the long wire to the remote ground renders the shield ineffective at high frequencies. It is a sobering fact that, while a device may control interference with respect to its own ground reference, it may have little or no effect at the equipment ground.[53,54] Because all these cord-connected devices divert additional 60 Hz and high-frequency noise currents into the safety ground system, they often aggravate the very problem they claim to solve. External, cord-connected filters, or those built into outlet strips, can serve to band-aid badly designed equipment. As shown in Fig. 32-24 (Section 32.4.2), some equipment is sensitive because common-mode power line disturbances, especially at high frequencies, have essentially been invited in to invade the signal circuitry!

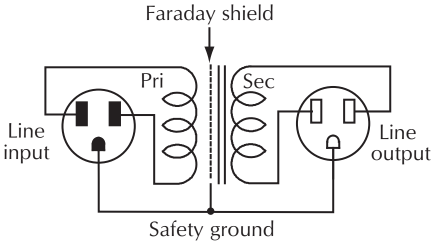


Figure 32-58. Power isolation transformer.

Second, the advertised noise attenuation figures for virtually all these power line devices are obtained in a most unrealistic way. Measurements are made with all equipment (generator, detector, and device under test) mounted to a large metal *ground plane*. Although the resulting specs are impressive, they simply don't apply to performance in real-world systems where ground connections are made with mere wires or conduit. However, these devices can be very effective when installed at the power service entrance or a subpanel, where all system safety grounds are bonded to a common reference point.[55] For thorough, accurate information about separately derived power distribution and its application to equipment racks, the author highly recommends reference 60.

Balanced power, more properly *symmetrical power*, is another seductively appealing concept shown in Fig. 32-59. If we assumed that each system box had neatly matched parasitic capacitances from each leg of the power line to its chassis ground, the resulting noise current flow into the safety ground system would be zero, the interchassis voltage would be zero, and the resulting system noise due to these currents would simply disappear! For example, if C_1 and C_2 had equal capacitance and the ac voltages across them were equal magnitude but opposite polarity, the net leakage current would indeed be zero. However, for the overwhelming majority of equipment, these capacitances are not equal or even close. In many cases, one is several times as large as the other—it's just a reality of power transformer construction. Even if the equipment involved has two-prong ac power connections, actual noise reduction will likely be less than 10 dB and rarely exceed 15 dB. And it's unlikely that equipment manufacturers will ever pay the premium to match transformer parasitic capacitances or use precision capacitors in power line EMI filters. If the equipment involved has three-prong (grounding) ac power connections, the leakage current reduction, if any, provided by symmetrical power will pale by comparison to the magnetically induced voltage differences described in Section 32.3.4. In fact, many of the benefits attributed to symmetrical power may result from simply plugging all system equipment into the same outlet strip or dedicated branch circuit—which is always a good idea.

A GFCI (ground-fault circuit interrupter) works by sensing the difference in current between the hot and neutral connections at an outlet. This difference represents current from the hot conductor that is not returning via neutral. The worst-case scenario assumes that the missing current is flowing through a person. When the difference current reaches 4–7 mA—producing a very unpleasant but non-life-threatening shock—an internal circuit breaker removes power in a fraction of a second. Some power conditioners feature a ground lift switch, touted to eliminate ground loop problems, at their outputs. The National Electrical Code requires that all balanced power units have GFCI-protected outputs (among other restrictions on the use of balanced power). Although safe, ground lifting makes a GFCI-protected circuit prone to nuisance trips. For example, consider the system hook-up shown in Fig. 32-60.

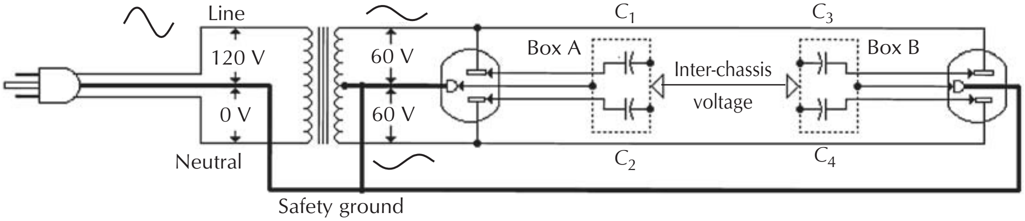


Figure 32-59. Balanced power hopes to cancel ground currents.

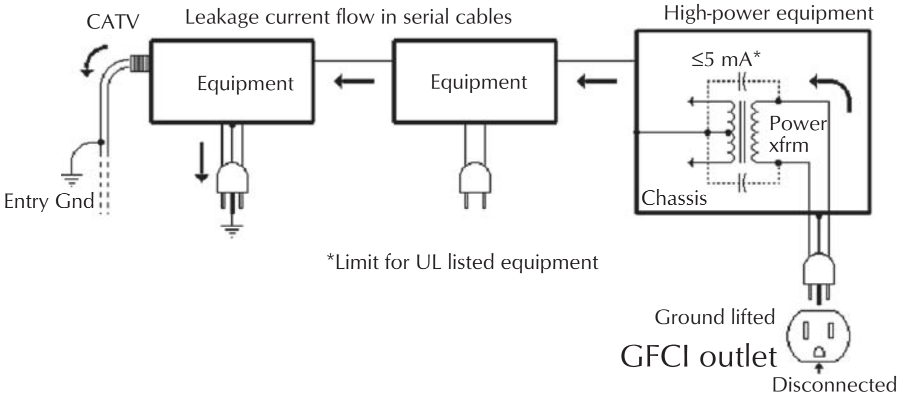


Figure 32-60. Common scenario to produce nuisance trips of GFCI in power conditioner.

For equipment having a grounding (three-conductor) power cord, UL listing requires that its leakage current be no more than 5 mA. Normally, this current would flow through the safety ground path back to neutral and would not trip a GFCI that has an intact safety ground connection. However, if the safety ground is lifted and the equipment is connected to other system equipment via signal cables, the leakage current will flow in these cables to reach ground, and ultimately neutral. Because the current is *not* returning via the equipment's own power cord, the GFCI considers it hazardous and may trip, since 5 mA is within its trip range. If multiple pieces of equipment are plugged into a single GFCI-protected circuit, the cumulative leakage currents can easily become high enough to trip the GFCI. This problem severely limits the ability of the GFCI/ground-lift combo to solve ground loop problems—even when balanced power partially cancels leakage currents.

32.7.3 Surge Protection

Haphazard placement of common *surge protectors* can actually result in damage to interface hardware if the devices are powered from different branch circuits.[56] As shown in Fig. 32-61, very high voltages can occur should there be an actual surge. The example shows a common protective device using three metal-oxide varistors, usually called MOVs, which limit voltage to about 600 V_{peak} under very high-current surge conditions.

For protection against lightning-induced power line surges, this author strongly recommends that MOV protective devices, if used at all, be installed only at the main service entry. At subpanels or on branch circuits to protect individual groups of equipment, use series-mode suppressors, such as those by Surge-X, that do not dump surge energy into the safety ground system, creating noise and dangerous voltage differences.[57,58]

32.7.4 Exotic Audio Cables

In the broadest general meaning of the word, every cable is a transmission line. However, the behavior of audio cables less than a few thousand feet long can be fully and rigorously described without transmission line theory. But this theory is often used as a starting point for pseudotechnical arguments that defy all known laws of physics and culminate in outrageous performance claims for audio cables. By some estimates, these specialty cables are now about a $200 million per year business.

Beware of cable mysticism! There is nothing unexplainable about audible differences among cables. For example, it is well known that the physical design of an unbalanced cable affects common-impedance coupling at ultrasonic and radio frequencies. Even very low levels of this interference can cause audible spectral contamination in downstream amplifiers.[59] Of course, the real solution is to prevent common-impedance coupling in the first place with a ground isolator, instead of agonizing over which exotic cable makes the most pleasing subtle improvement. Expensive and exotic cables, even if double or triple shielded, made of 100% pure unobtainium, and hand braided by Peruvian virgins, will have NO significant effect on hum and buzz problems! As discussed in Section 32.5.4, shielding is usually a trivial issue compared to common-impedance coupling in unbalanced interfaces. It's interesting to note that some designer cables selling for $500/meter pair have no overall shield at all—ground and signal wires are simply woven together.

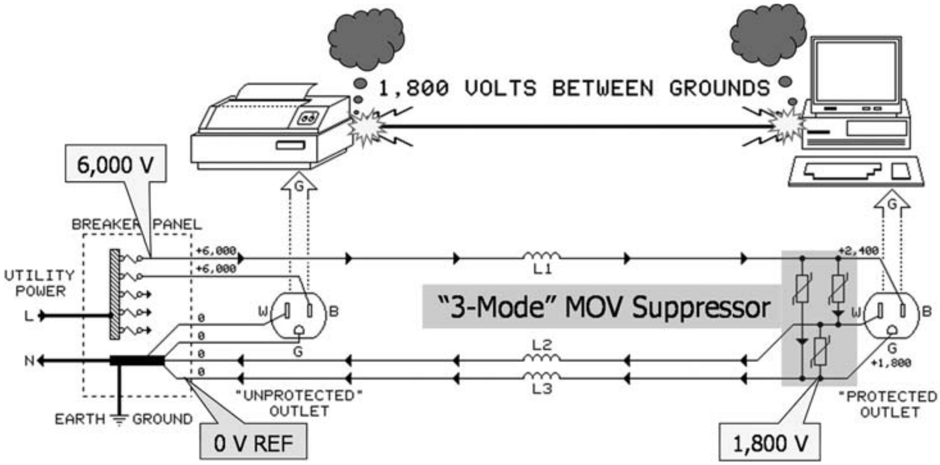


Figure 32-61. Surge protection can actually invite equipment damage.

Some exotic audio cables have very high capacitance and can seriously degrade high-frequency response, especially if cables are long and/or a consumer device drives it. For demanding high-performance applications, consider a low-capacitance, low-shield-resistance cable such as Belden #8241F. Its 17 pF/ft capacitance allows driving a 200 ft run from a typical 1 kΩ consumer output while maintaining a –3 dB bandwidth of 50 kHz. And its low 2.6 mΩ/ft shield resistance, equivalent to #14 gauge wire, minimizes common-impedance coupling. It's also quite flexible and available in many colors.

References

1. Grover, F. W., *Inductance Calculations*, New York: Van Nostrand, 1944, p. 35, eq. (7).
2. Ott, H., *Noise Reduction Techniques in Electronic Systems*, Second Edition, John Wiley & Sons, 1988, p. 116.
3. Davis, D. and C., *Sound System Engineering,* Second Edition, Focal Press, 1987, p. 382.
4. *Audio Test Procedures*, ATP1.1 through ATP1.4 and Appendix ATP.C, British Broadcasting Corporation, circa 1978–1980.
5. Hess, R., *Voltage Transmission for Audio Systems*, Audio Engineering Society 67th Convention, 1980, Preprint #1708.
6. Ott, H., op. cit., p. 31.
7. Gerke, D., and Kimmel, W., *Kimmel Gerke Bullets*, Kimmel Gerke Associates, Summer 2000, p. 3.
8. Ott, H., op. cit., pp. 47–49.
9. Morrison, R., and Lewis, W., *Grounding and Shielding in Facilities*, New York: Wiley-Interscience, 1990, pp. 144–145.
10. Gerke, D., and Kimmel, W., *The Mysteries of Grounding*, IEEE EMC Society Newsletter, Issue 178, Summer 1998, pp. 13–14.
11. Morrison, R., and Lewis, W., op. cit., pp. 47–49.
12. Whitlock, B., "Surge Protection: The Enemy Within," *Sound & Video Contractor*, July 1999, pp. 94–95.
13. Whitlock, B., "Ground Rods: Magical Noise Sinks?" *Sound & Video Contractor*, November 2000, pp. 102–103.
14. Whitlock, B., "Neutral Ground-Finding Swaps in Power Wiring," *Sound & Video Contractor*, June 2001.
15. Denny, H., *Grounding for the Control of EMI*, Don White Consultants, Inc., Gainesville, VA, 1989, pp. 9.8–9.9.
16. Hiser, S., *1997 Electrocutions Associated with the Use of Consumer Products*, U. S. Consumer Products Safety Commission, 2000, www.cpsc.gov
17. Mah, J. et al., *1997 Residential Fire Loss Estimates*, U.S. Consumer Products Safety Commission, 2000, www.cpsc.gov.
18. Denny, H., op. cit., pp. 7.2–7.3.

19. Ott, H., op. cit., p. 81.
20. Ott, H., op. cit., p. 106.
21. Morrison, R., *Solving Interference Problems in Electronics*, First Edition, John Wiley & Sons, 1996, p. 35.
22. Morrison, R. and Lewis, W., op. cit., pp. 168–169.
23. *Reference Data for Radio Engineers*, Fifth Edition, Howard W. Sams, 1972, Section 11-1.
24. Whitlock, B., "Balanced Lines in Audio Systems: Fact, Fiction, and Transformers," *J. Audio Eng. Soc.*, June 1995, pp. 454–464.
25. Whitlock, B., *A New Balanced Input Circuit for Maximum Common-Mode Rejection in Real-World Environments*, Audio Engineering Society 101st Convention, 1996, Preprint #4372.
26. Ott, H., op. cit., p. 105.
27. Morrison, R., *Noise and Other Interfering Signals*, John Wiley & Sons, 1992, p. 55.
28. Whitlock, B., "Balanced Lines in Audio Systems: Fact, Fiction, and Transformers," *J. Audio Eng. Soc.,* June 1995, pp. 460–462.
29. Muncy, N., "Noise Susceptibility in Analog and Digital Signal Processing Systems," *J. Audio Eng. Soc.*, June 1995, pp. 441–442.
30. Muncy, N., op. cit., p. 436.
31. Perkins, C., "Automated Test and Measurement of Common Impedance Coupling in Audio System Shield Terminations," *J. Audio Eng. Soc.*, June 1995, pp. 488–497.
32. Windt, J., "An Easily Implemented Procedure for Identifying Potential Electromagnetic Compatibility Problems in New Equipment and Existing Systems: The Hummer Test," *J. Audio Eng. Soc.*, June 1995, pp. 484–487.
33. Whitlock, B., *Hum & Buzz in Unbalanced Interconnect Systems*, Application Note AN004, Jensen Transformers, Inc., 1996.
34. Morrison, R., *Noise and Other Interfering Signals*, John Wiley & Sons, 1992, pp. 136–137.
35. Muncy, N., op. cit., p. 436.
36. Ott, H., op. cit., p. 134.
37. Jensen, D., and Sokolich, G., *Spectral Contamination Measurement*, Audio Engineering Society 85th Convention, 1988, Preprint #2725.
38. Whitlock, B., "Understanding and Controlling RF Interference," *Sound & Video Contractor*, February 1999, pp. 78–80.
39. Fielder, L., "Dynamic Range Issues in the Modern Digital Audio Environment," *J. Audio Eng. Soc.,* May 1995, pp. 322–339.
40. Pease, R., *Troubleshooting Analog Circuits*, Butterworth-Heinemann, 1991.
41. Whitlock, B., *Interconnection of Balanced and Unbalanced Equipment*, Application Note AN003, Jensen Transformers, Inc., 1995.
42. Whitlock, B., *Understanding and Controlling RF Interference*, Sound & Video Contractor, February 1999, pp. 78–80.
43. Fair-Rite Products, One Commercial Row, Wallkill, NY 12589, www.fair-rite.com.
44. Steward, 1200 East 36th Street, Chattanooga, TN 37401, www.steward.com.
45. Perkins, C., "To Hum or Not to Hum," *Sound & Video Contractor*, March 1986, p. 41.
46. Leach, W., "The Differential Time-Delay Distortion and Differential Phase-Shift Distortion as Measures of Phase Linearity*," Journal of the Audio Engineering Society*, September 1989, pp. 709–715.
47. Jensen, D., and Sokolich, G., *Spectral Contamination Measurement*, Audio Engineering Society 85th Convention, 1988, Preprint #2725.
48. Jensen, D., *High-Frequency Phase-Response Specifications—Useful or Misleading?* Audio Engineering Society 81st Convention, 1986, Preprint #2398.
49. Whitlock, B., *Hum & Buzz in Unbalanced Interconnect Systems*, Application Note AN004, Jensen Transformers, Inc., 1996.
50. Morrison, R., and Lewis, W., op. cit., pp. 95–98.
51. Whitlock, B., "Ground Rods: Magical Noise Sinks*?" Sound & Video Contractor*, November 2000, pp. 102–103.
52. Morrison, R., *Solving Interference Problems in Electronics*, First Edition, John Wiley & Sons, 1996, p. 1.
53. Morrison, R., and Lewis, W., op. cit., pp. 106, 157–158, 177.
54. Morrison, R., *Solving Interference Problems in Electronics*, First Edition, John Wiley & Sons, 1996, p. 9.

55. Morrison, R., *Solving Interference Problems in Electronics*, First Edition, John Wiley & Sons, 1996, pp. 140–146.
56. Whitlock, B., "Surge Protection: The Enemy Within," *Sound & Video Contractor*, July 1999, pp. 94–95.
57. Harford, J., "Will My Surge Suppressor Work?" *Sound & Video Contractor*, September 1997.
58. Morrison, R., and Lewis, W., op. cit., pp. 159-160.
59. Jensen, D., and Sokolich, G., op. cit.
60. Middle-Atlantic Products, *Integrating Electronic Equipment and Power into Rack Enclosures—Optimized Power Distribution and Grounding for Audio, Video and Electronic Systems*, 2008, free white paper download at www.middleatlantic.com/power.htm.

Suggested Reading

Henry Ott, *Noise Reduction Techniques in Electronic Systems*, Second Edition, John Wiley & Sons, 1988.

Ralph Morrison and Warren Lewis, *Grounding and Shielding in Facilities*, John Wiley & Sons, 1990.

Ralph Morrison, *Grounding and Shielding Techniques in Instrumentation*, Third Edition, John Wiley & Sons, 1986.

Trademarks

ISO-MAX® is a registered trademark of Jensen Transformers, Inc.

InGenius® is a registered trademark of THAT Corporation.

Romex® is a registered trademark of General Cable Industries, Inc.

This page intentionally left blank

Chapter 33

System Gain Structure

by Pat Brown

33.1 Introduction 1223
  33.1.1 Interfaces 1223
33.2 Audio Waveform Fundamentals 1224
33.3 Gain Structure 1225
33.4 System Gain Structure 1227
33.5 The Unity Method 1228
33.6 An Optimized Method 1229
33.7 Setting the Amplifier Sensitivity 1230
33.8 Power Ratings 1231
  33.8.1 Amplifier Power Ratings 1231
  33.8.2 Loudspeaker Power Ratings 1231
33.9 Conclusion 1232
Reference 1232

This page intentionally left blank

33.1 Introduction

This chapter is devoted to the explanation and establishment of the proper gain structure of the sound reinforcement system. It has been the author's experience that sound systems are rarely producing the optimum performance that would be indicated by the specification sheets of the individual components. Tangible improvements in performance can often be achieved by some simple adjustments of level controls.

Most technical subjects can be best explained using ideal relationships, and this one is no exception. The real world always falls short of the ideal case, but the ideal can present a model and goal for our efforts. It is the responsibility of the sound practitioner to form an understanding of the trade-offs and apparent contradictions through experience and endless hours in the field. What follows is only an introduction that will benefit those who supplement it with lab and field work.

33.1.1 Interfaces

An interface exists when two components are to be interconnected for the purpose of transferring a signal. One component will be the source (sending) device and the other the load (receiving) device for the electrical signal. At least three major topologies exist for interconnecting devices, the major difference being which electrical parameter of the signal that the interface optimizes the passage of—i.e., voltage, current, or power. This is primarily a function of the ratio between the source impedance and load impedance. At this point we will make our first simplification by assuming that the impedance of these devices is purely resistive with no appreciable reactive component. This is actually a pretty accurate assumption for most electronic components in the signal processing chain.

33.1.1.1 The Matched Interface

A matched interface means that the source and load impedances are equal. This topology has some admirable attributes:

1. Power transfer is maximized.
2. Reflections from load-to-source are eliminated.

Impedance matching is required when the electrical wavelengths of the audio signal are shorter than the interconnecting cable. Examples include antenna circuits, digital interfaces, and long analog telephone lines. A drawback of this interface is that power transfer is optimized at the expense of voltage transfer, and therefore the source device might be called on to source appreciable current. It is also more difficult to split a signal from one output to multiple inputs, as this upsets the impedance match. A component that is operated into a matched impedance is said to be *terminated*. While the telephone company must use the matched interface due to their electrically long lines, the audio industry departed from the practice many years ago in favor of the voltage-optimized interface for analog interconnects.

Fig. 33-1 shows a matched interface. It is important to note that the selection of 600 Ω as the source and load impedance is arbitrary. It is the impedance ratio that is of importance, not the actual value used.

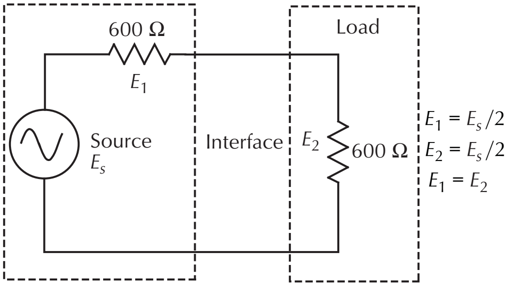


Figure 33-1. Matched interface.

33.1.1.2 The Constant Voltage Interface

Most analog sound system components are designed to operate under *constant voltage* conditions. This means that the input impedance of the driven device is at least ten times higher than the output impedance of the source device. This mode of operation assures that output voltage of the driving device is relatively independent of the presence of the driven device—hence the term constant voltage, Fig. 33-2. Constant voltage interfaces can be used in analog audio systems since the typical cable length is far shorter than the electrical wavelength of the signal propagating down the cable.

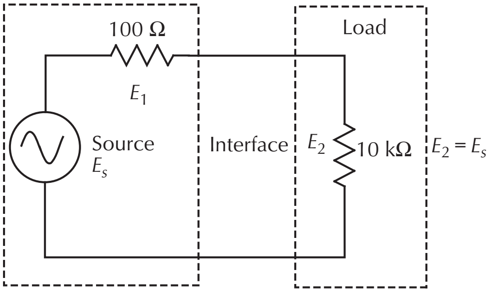


Figure 33-2. Constant voltage interface.

This makes such lines immune to the detrimental effects of reflections and standing waves. Radio, digital, and telephone engineers are not so fortunate, and impedance matching is required at component interfaces. Constant voltage (sometimes called bridged) interfaces are inherently simpler than their impedance matched cousins. Advantages include the ability for a single output to drive multiple high-impedance inputs (in parallel) without loss of signal or degradation. Also, the constant voltage interface does not require that manufacturers standardize their input and output impedances. As long as the output impedance is low (typically less than 1000 Ω) and the input impedance is high (typically greater than 10 kΩ) then the two devices are compatible. In practice most output impedances are fairly low (<100 Ω), allowing a single low-impedance output to drive several high-impedance inputs, Fig. 33-3.

If the source impedance is large when compared to the load, then a constant current interface is formed. In this topology, the current from the source is determined by the source impedance and is independent of the load impedance. Constant current interfaces are not often used to interface electronic components, and are usually reserved for specialized applications, such as the construction of impedance meters. We will not consider this interface further in this chapter.

33.2 Audio Waveform Fundamentals

In a sound reinforcement system, *program sources* provide information that is to be reinforced and presented to a listener. This information can originate in the form of an acoustical wave (acoustic musical instruments or human voices) or an electrical wave (electronic instruments or storage media such as compact disc). In either case, the waveform must be in the electromagnetic domain prior to being presented to the sound system. Acoustic signals must be converted into electromagnetic signals with an appropriate transducer such as a microphone or accelerometer. We will refer to electromagnetic waves within the bandwidth of the human auditory system as *audio* waveforms. Typical audio waveforms are quite complex and are continuously changing in value over time. This makes it difficult to describe them numerically. Several parameters can be used to describe the characteristics of an audio waveform. These include the following:

Peak-to-Peak Voltage. The number of volts between the largest positive and largest negative peak of the waveform.

Peak Voltage. The highest peak of the waveform, regardless of whether it is positive or negative. For a waveform with amplitude symmetry, it will be one-half the peak-to-peak voltage.

Average Voltage. The average of all ± amplitude values of the waveform.

Root-Mean-Square (rms) Voltage. Sometimes called the *effective* value of the waveform, rms describes the ac voltage in terms of the equivalent dc voltage that would produce the same amount of heat into a resistive load. Rms is useful because it indicates the heating value of the waveform. The rms level of a complex audio waveform is also related to its perceived loudness if it is used to drive a loudspeaker. For a sine wave, the rms voltage is 0.707 times the peak voltage. Complex waveforms also have an rms voltage, but finding it requires integration of the waveform over time. The peak-to-rms ratio of a waveform is called its *crest* factor. Crest factors must be described in terms of a finite span of time. A span of 50 ms correlates well with the integration time of the human hearing system, but other values can be used.

Fig. 33-4 shows a sine wave and speech waveform on a common plot of amplitude versus time. An oscilloscope provides this representation of the data, as does a wave editor application for a personal computer. A

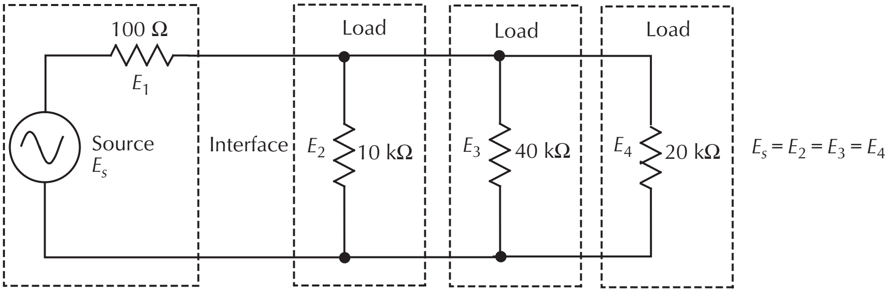


Figure 33-3. Constant voltage bridged interface driving multiple inputs.

simple analog voltmeter can measure the rms value of a sine wave. Complex waveforms require more sophisticated instruments to yield their effective value.

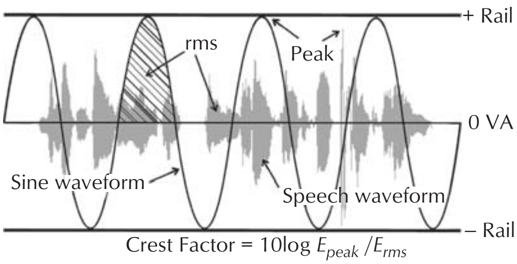


Figure 33-4. Crest factors for sine waves and speech waves.

The peak voltage of the waveform must pass through the sound system component without being clipped. It is the parameter of interest when establishing the gain structure of the system. The crest factor of the signal determines the energy content and therefore the power produced by the amplifier and delivered to the loudspeaker. It is of interest when considering the heat that the loudspeaker must dissipate. Additionally, the rms voltage is the signal parameter that relates most closely to the loudness of the signal as perceived by a human listener.

Since the goal of an audio system is to reproduce appropriate waveforms for a given application, these waveform principles have universal application for all parts of the sound system.

33.3 Gain Structure

The following are general terms of gain structure beginning with how it applies to an individual piece of electronic equipment. It matters little whether the equipment is a mixer, equalizer, amplifier, or other active system component. By active we mean that the device has a power supply for its internal active circuitry. This can be as simple as an internal battery or two, or as complex as an internal or external ac line-powered supply. The power supply voltages establish the maximum amplitude that a waveform can take on as it passes through the component, Fig. 33-4. In audio equipment, most power supplies form a bipolar set of rails—a fixed positive and negative zero frequency (dc) voltage that the waveform is developed between. The value of the rail voltage determines the peak amplitude that the waveform can take on. Exceeding this peak value will cause deformation of the wave, commonly known as *clipping*. We will proceed under the assumption that the rails are fixed, and indeed they are for most signal processing devices. Some power amplifier topologies use multivalued or fluctuating rails. The principles are the same, but we will not consider such devices here.

Under a no input signal condition, all audio components will still emit a residual output signal. Thermal noise is generated at the molecular level and is present at the output of all system components whether active or passive. The level of the thermal noise determines the noise floor of the component. In practice, other factors can also make a contribution to the residual noise of an electronic device. An undersized power transformer or poor shielding can elevate some frequencies above the broadband thermal noise floor, Fig. 33-5. Equipment designers try to minimize thermal noise by component selection and careful design, but it can never be eliminated. We must accept the fact that it exists. Part of the reason for establishing the proper gain structure of a system is to render the effects of thermal noise insignificant. The thermal noise floor is affected by the settings of the component's level controls. While a low noise floor can be achieved with all controls set at minimum, this is not realistic, as we cannot operate it that way. The controls should be set at a point appropriate for the operation of the device. A good starting point is a setting that produces the same voltage at the device output that is present at the device input, often called *unity* by audio practitioners. Level controls placed at their 0 dB setting generally produce this condition, and represent a good starting point for setting up a system.

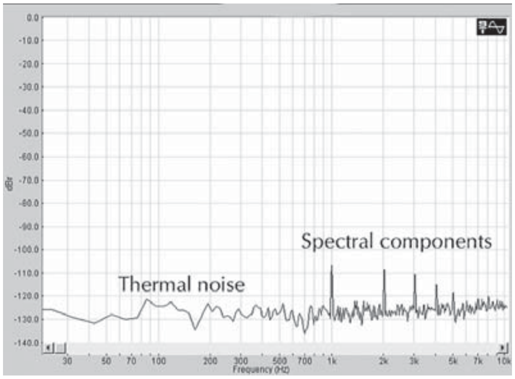


Figure 33-5. Thermal and spectral noise.

Knowledge of the supply rails and noise floor establishes a dynamic range for the device—the difference in level between the highest possible undistorted peak and the lowest level that the signal can take on without being buried in the noise. The dynamic range is what *can* happen when a signal is passed through a device. It is a range of possible values that the waveform can take on. The possibilities are infinite (within a device's dynamic range) for an analog component and finite for digital components, since digital signals are made up of discrete samples that must be quantized to fixed steps. Fig. 33-6 shows a 1 kHz sine wave driving a component to just below clipping. The level difference between clipping and the noise floor describes the component's dynamic range.

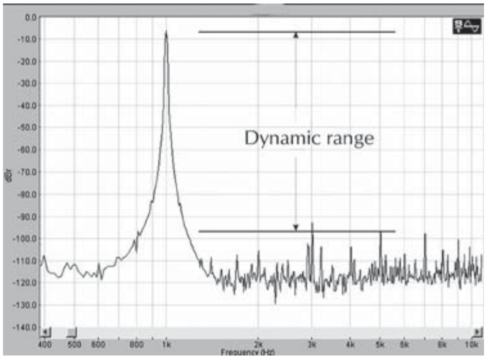


Figure 33-6. Dynamic range.

An example is in order. Let us consider a line level audio signal processor. We can pick any rail voltage that we like, since no universally recognized standard exists. A rail voltage of +17.5 Vdc (and –17.5 Vdc for the negative rail) will allow a peak voltage of 17.5 V to be realized at the device output. It is customary to express this voltage in terms of the rms value of the largest sine wave that the device (peak –3 dB) can produce with an acceptable amount of distortion. An oscilloscope allows observation of the wave and any deformation due to clipping. This rms voltage becomes the maximum output voltage, and when expressed in decibels becomes the component's maximum output level. This level is expressed in dBm (dB ref. 0.001 W) for impedance matched interfaces (note that knowledge of the circuit impedance is required) or dBV (dB ref. 1 V) or dBu (dB ref. 0.775 V) for constant voltage interfaces (assuming the bridged impedance condition is maintained).

$$
\begin{aligned}
L_{out} &= 20\log(17.5) - 3 \\
&= 21.8 \text{ dB dBV} \\
&= 20\log\left(\frac{17.5}{0.775}\right) - 3 \\
&= 24 \text{ dBu}
\end{aligned}
$$

Assume that the thermal noise measured at the device output is about 200 μVrms, as measured using an rms broadband voltmeter. Expressed as a level in dBV, the thermal noise floor becomes:

$$
\begin{aligned}
L_{noise} &= 20\log(0.0002) \\
&= -74 \text{ dBV} \\
&= 20\log\left(\frac{0.0002}{0.775}\right) \\
&= -71.8 \text{ dBu}
\end{aligned}
$$

This audio component thus has a dynamic range on the order of 100 dB, a very good figure, and one that is typical for a well-designed piece of audio equipment, whether analog or digital.

With the dynamic range established, it is still necessary to use it effectively. If a weak signal is fed to the component, it may fall far short of the clipping point established by the power supply voltages, placing it unnecessarily close to the component's noise floor. This will produce a poor signal-to-noise ratio, SNR, even in a component that has a wide dynamic range, Fig. 33-7.

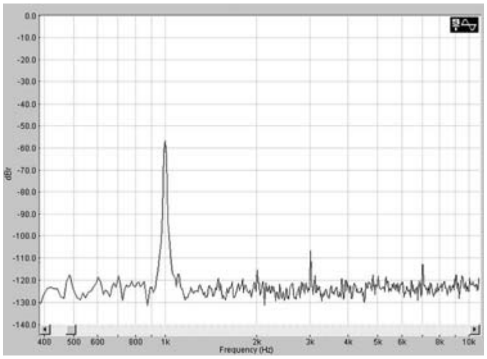


Figure 33-7. Poor SNR.

If the input level control is increased beyond unity, the thermal noise will likely increase with the signal voltage, and no increase in SNR is realized. Increasing the drive (source) voltage will improve the SNR, assuming that the sending device has a noise floor lower than the driven device. In some cases, an additional gain stage may be required, as with microphones and phonograph cartridges.

If too strong a signal is fed to the component, the highest amplitude parts of the waveform may not fit within the constraints of the power supply voltages and may drive the component into a nonlinear mode of operation (clipping). This may yield an excellent signal-to-noise performance, but a distorted output signal rich in harmonic distortion, Fig. 33-8. Gain structure, from a component perspective, is passing the signal through at optimum amplitude—not too strong and not too weak. As such, a system component can be overdriven, underdriven, or optimally driven by a signal source. It is important to note that the SNR of the program source is often the determining factor for the SNR of the entire system, since it can only be improved with very specialized signal processing that is not found in typical reinforcement systems. The old adage “garbage in, garbage out” certainly applies. The *SNR* will be degraded as it passes through other system components, which is why care must be taken to properly calibrate each stage of the system.

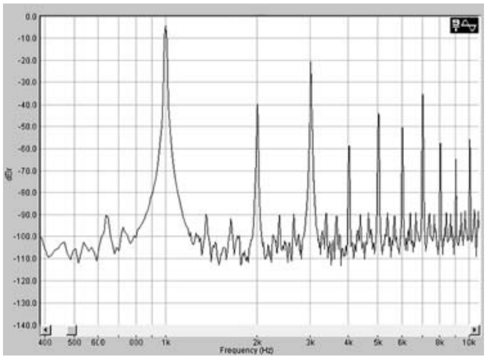


Figure 33-8. Overdriven signal characteristics.

33.4 System Gain Structure

Audio components have been evolving since early in the last century. In the process the dynamic ranges of system components have become similar, and in many cases approach the theoretical limits dictated by nature. While overall dynamic range may be similar, the clipping levels and noise floors are not identical between manufacturers or even within product lines. While we won’t go into all of the reasons for this, it is unfortunate that at least clipping levels aren’t standardized within the sound reinforcement industry. As such, it is possible for a system component to be operating optimally within its own dynamic range, and yet be overdriving or underdriving the next component. This reality forces us to consider gain structure from system perspective.

Before discussing the gain structure of a sound system, it is necessary to consider a method for determining the internal gain structure of a system component. This can be done by introducing a stimulus to the component and observing its output signal. It is common practice for technicians to use a stable and repeatable waveform for calibrating the signal-processing chain. A sinusoidal waveform, commonly called a sine wave, is such a waveform. The sine wave is a single frequency tone that is easily generated, fed to an input, and observed at the output of each component in the chain. The previous graphs have shown sine waves displayed on a magnitude versus frequency plot. They resemble a vertical spike due to the narrow frequency bandwidth. An alternate and equally valid display is amplitude versus time. The oscilloscope displays the amplitude of the waveform as a function of time. More advanced models will even provide some statistics, such as peak voltage, rms voltage, frequency, crest factor, level, etc. Let us proceed. A 1000 Hz sine wave is developed across the input terminals of one channel of the mixer. An amplitude is chosen that does not overdrive the input, and that is of sufficient level to drive the mixer to its full undistorted output voltage with all level controls in the signal path set at unity. For microphone inputs, about 0.1 Vrms (–20 dB ref. 1 V) is generally sufficient. As much as 1 Vrms might be required for a line level input. The level controls of the mixer are set as follows:

- Master at unity or 0 dB.
- Channel at unity or 0 dB.
- Trim at unity or 0 dB.

Under these conditions, the voltage amplitude of the output signal should be the same as the voltage amplitude of the input signal—an amplification factor of one or unity, and a gain of 0 dB.

The input voltage has been increased until the main meter of the mixer reads zero. We will speak in more detail about zero later, but for now we will assume that it indicates a voltage in the optimum operating range of the mixer’s overall dynamic range (typically –20 dB rel. clipping). Since program audio waveforms are constantly changing, this operating level allows some room for peaks in the audio waveform to pass undistorted. It is instructive at this point to measure the output voltage of the mixer at meter zero. Using the oscilloscope, either the peak or rms value of the waveform can be measured. It is traditional to measure the rms value, since it is readily measured with much less sophisticated voltmeters than the oscilloscope and correlates well with the loudness and heat production of the signal.

For historical reasons, a common voltage measured at a mixer output with the meter indicating zero is 1.23 Vrms, corresponding to an rms open circuit level of +4 dB ref. 0.775 V (+4 dBu). This level might be termed the *operating level* of the mixer. A volume indicating meter describes the waveform in a way that correlates with its loudness as perceived by human listeners. The indication of such a meter is in VU, or volume units. When sound systems used the impedance matched interface, this voltage was developed across the input impedance (usually 600 Ω) of the next input stage. Since voltage and impedance were known, the power equation could be used to calculate the output power of the mixer, which became

$$W = \frac{E^2}{R}$$

$$L_{out} = 10\log\frac{1.23^2}{600}$$

$$= 4\ \text{dBm}$$

The power transfer was relevant since the matched interface was used, optimizing the circuit for power transfer. One milliwatt provided a useful reference, as it falls in the middle of the range of power levels found in the sound system. A voltmeter calibrated to read zero at 0.775 V could directly indicate the circuit power level in dBm (assuming a 600 Ω matched impedance interface). This calibration would make the voltmeter a dBm meter when placed across a 600 Ω circuit. When dBm meters are used at other impedances, a correction factor is required. As the sound reinforcement industry migrated to the constant voltage interface, the 0.775 V reference lived on due to the proliferation of voltmeters so calibrated, and signals were then described in dB ref. 0.775 V or dBu. In modern systems, the term *level* is used to describe the field quantity of interest at a component interface, which is the signal voltage for the constant-voltage interface. This a good place to note that many modern mixers do not use the +4 dBu reference level for meter zero, so the reader is advised to consult the literature or perform a measurement. A more common meter zero level today is 0 dBV, or 1 Vrms.

Let us now advance the trim control (or the drive voltage) until the waveform becomes distorted when viewed on the scope. Some mixers have a clipping indicator to warn of this condition. When the waveform flattens on top, reduce the trim control until the waveform appears undistorted. Since mixers are made up of several stages, it is usually informative to move each the main fader, channel fader, and trim control until clipping is observed to assure that each stage is clipping simultaneously. This produces the maximum output voltage of the mixer at the mixer's output terminal. Using the scope or a voltmeter, measure the voltage of the waveform. Note that the clipping occurs on the peak of the waveform, yet it is standard practice to measure the rms value of the waveform and include it on the specification sheet. Ideally, this maximum output voltage is at least ten times the voltage measured at the meter zero indication, providing 20 dB of peak room above meter zero. The drive level (or trim control setting) should now be reduced to produce the meter zero operating level of the mixer.

We now have knowledge of the operating and clipping level of the mixer (e.g., +4 dBu and +24 dBu respectively). These values should be recorded in the system documentation. The noise floor of the mixer can be measured by muting the input signal and measuring the mixer's no signal output voltage, but this is of little interest in practice.

33.5 The Unity Method

Our mixer is now at an optimum operating level with good SNR and 20 dB of peak room. The signal from the mixer is fed to the next component in the chain. If the component has input and/or output level controls, they are adjusted to produce the same level of the mixer at that component's output terminal. In like manner the signal is fed through subsequent signal processors, and the mixer's level eventually ends up at the input of the power amplifier, whose input sensitivity control is set for the desired output voltage (i.e., the target playback level of the system). As the amplifier's voltage is impressed across the loudspeaker load, the amplifier supplies current flow as determined by the impedance of the loudspeaker. Power will flow, but the signal level is a linear function of the applied voltage over the useful operating range of the amplifier. So, the output voltage is the parameter of interest in a properly configured amplifier-to-loudspeaker interface under normal operating conditions. Fig. 33-9 shows such a processing chain for a mixer with a 0 dBV meter zero. The unity amplification method has a number of advantages, which include:

1. Ease of calibration.
2. Fast implementation.
3. Easy substitution of components.

Unfortunately, there are some drawbacks to this approach, mostly due to the nonstandardization of clipping levels between product lines and manufacturers. A mixer operating at 0 dBV that clips at +20 dBV will have 20 dB of operating peak room for transient peaks. If the component after the mixer clips at +18 dBV, that component will only have 18 dB of operating peak room. In this case an undistorted full-scale waveform from the mixer would cause clipping in the next component. The

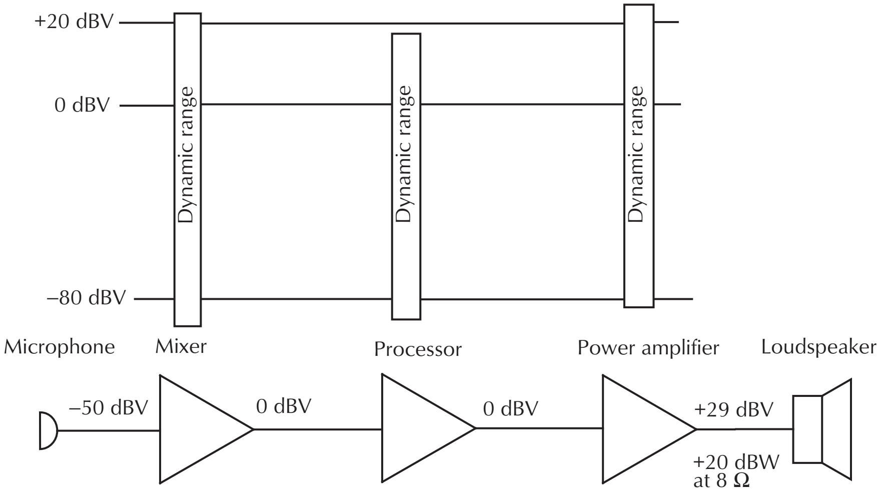


Figure 33-9. The processing chain.

mixer could be turned down a bit if the overload is not severe. If the level mismatch is more than a few dB then a different solution may be required. It should be pointed out that this condition is not as prevalent as it once was, as many postmixer product manufacturers have modified their products to handle the higher output voltages produced by modern mixing consoles.

33.6 An Optimized Method

The drawbacks of the unity amplification method can be overcome with an optimized method of establishing the gain structure of the system. While the unity method establishes a consistent operating voltage from component to component, the optimized method sets each device to clip simultaneously, regardless of the actual signal level. A mixer outputting +24 dBV and an equalizer outputting +20 dBV are both set to reach their clipping point simultaneously. This method often requires the insertion of a resistive attenuator between the mixer and equalizer, allowing the mixer to output its maximum voltage and still not over-drive the equalizer.

To optimize the system gain structure, feed a sine wave to the mixer in the same manner previously described, but this time advance the trim control until clipping occurs at the mixer output. All power amplifiers should be off or fully attenuated. The clipping can be determined with the aid of an oscilloscope or a spectrum analyzer capable of handling at least +30 dB ref. 1 V (+30 dBV). With the mixer set just short of clipping, connect the output of the mixer to the input of the next component (i.e., an equalizer). Set all controls on the equalizer at their unity setting. Move the clipping indicator (scope) to the output of the equalizer and note whether the waveform is clipped. If it is not, the equalizer is capable of passing the full output voltage of the mixer. If the equalizer is clipping, first try reducing the setting of its input level control. This often doesn't work since the stage being overdriven likely precedes the level control stage. Some manufacturers design their equipment to handle higher input levels than they can output, and the input level control may indeed eliminate the overdrive condition. If not, an attenuator is placed between the mixer and equalizer and adjusted to produce an undistorted waveform from the equalizer. The same procedure is repeated for each subsequent piece of equipment in the processing chain. Compressor/limiters should be set at their highest threshold setting and lowest compression ratio. Crossover networks require either:

1. Selection of a sine wave within the pass band of the output being tested, or
2. Readjust the crossover frequency control to allow the output being tested to pass the test signal. Remember to restore the setting before turning on the amplifier!

Resistive pads for the "too hot" source are available commercially. If a component is being *underdriven* by the previous one (i.e., the full output voltage of the mixer is insufficient to clip the equalizer), it may be advantageous to increase the underdriven component's input level until just short of clipping. This will provide a stronger drive voltage to the *next* component (and possibly an improved signal to noise).

The advantages of the optimized method include:

1. The ability to mix at meter zero without the potential to clip a component farther down the processing chain.
2. Optimized SNR in all components in the chain.
3. The meter mixer can now serve as an accurate indicator for all subsequent components, since all clip simultaneously.

As with all things audio, the optimized method is not without drawbacks. These include:

1. It requires more time and expertise to set up.
2. It requires a method to determine device clipping (scope or spectrum analyzer).
3. It requires the purchase of or construction of pads.
4. It makes component substitution more difficult, since the replacement component may have a different clipping level than the defective one.

A pad of 5–15 dB may be necessary for a professional mixer driving a consumer recorder.

Fig. 33-10 shows a system whose gain structure has been optimized in this manner. The benefit of either method is that a healthy drive voltage with a good SNR is delivered to the power amplifier input. Both methods also assure that any digital components in the processing chain are being driven with a voltage high enough to produce optimized A/D conversion.

33.7 Setting the Amplifier Sensitivity

Ideally the amplifier's input stage should handle the full output level of the preceding device without clipping. It is possible for the amplifier input circuit to clip prior to its output stage. This can be tested by setting the amplifier attenuator at a very low level and observing the output waveform of the amplifier when driven with the full undistorted signal level of the preceding device. If clipping is apparent at the amplifier output at a low attenuator setting, the input stage is being overdriven. Insertion of a pad or a reduction in drive voltage will be required.

If an active crossover is in the signal chain, its proper settings should be established prior to switching on and setting the amplifier input sensitivity. These settings are best obtained from the loudspeaker manufacturer.

The amplifier could be calibrated in the same manner as the other components—simply adjust its input sensitivity (volume) control to produce an output signal just short of clipping. Since this may be too loud, it is better to use a broadband program source (pink noise or music) and adjust the amplifier for the desired L_P at the listener position. The procedure is as follows.

With the program material being input to the mixer, the mixer is set to produce a zero meter indication as previously described. Note that this assumes a VI meter and not a peak program meter. The input attenuator of the power amplifier is then advanced until:

1. The desired acoustic level is reached in the audience, or
2. The amplifier begins to indicate clipping.

In either case, don't turn it up any higher. The gain structure is complete and the system is producing its maximum undistorted L_P. The technician can now proceed to fine tuning of the crossover network and equalizer to finish the system calibration.

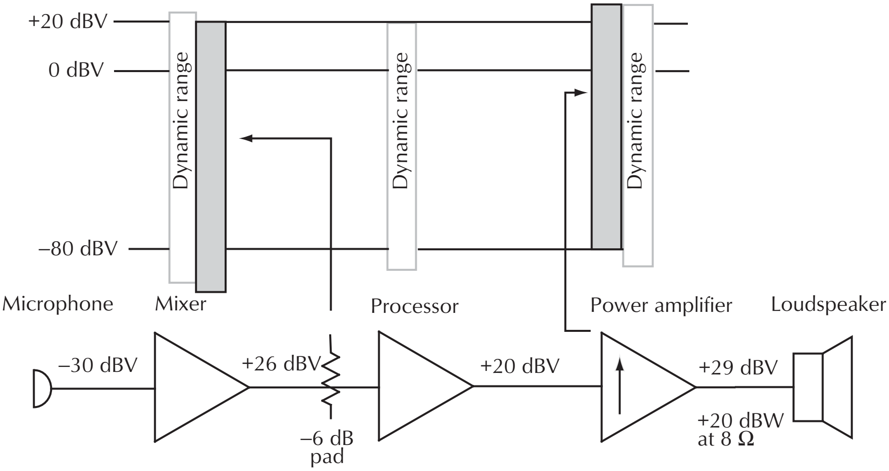


Figure 33-10. Optimized system gain structure.

33.8 Power Ratings

It is important to note that the amplifier's wattage rating must be *appropriate* for the loudspeaker. The loudspeaker must not be driven beyond its ability to dissipate heat buildup or beyond its limits of mechanical excursion. In most systems, heat will be a bigger problem than excursion, so the rms value of the amplifier's output voltage waveform must be managed. Since the continuous power (based on the rms voltage) delivered to the loudspeaker is largely a function of the type (crest factor) of the program material, selection can be complicated.

33.8.1 Amplifier Power Ratings

The power ratings of power amplifiers and loudspeakers have little in common. The amplifier is usually rated in accordance with the maximum continuous power that it can deliver reliably with sinusoidal input for a specified span of time into a specified load impedance. This yields a large number for the amplifier power rating (due to the high rms voltage of the sine wave), and most likely a wattage that the amplifier will never be called on to deliver, since the signals that we present to audiences usually bear little resemblance to sine waves. Even so, this rating can be useful for amplifier comparisons and selection. Just remember, you won't get that much rms voltage across the loudspeaker with real-world program material.

33.8.2 Loudspeaker Power Ratings

The loudspeaker's continuous power rating describes the loudspeaker's ability to dissipate heat on a continuous basis. A meaningful rating must state at a minimum:

1. The type and crest factor of the signal used.
2. The bandwidth of the signal.
3. The time duration of the test.
4. The rms voltage of the signal.
5. The impedance of the loudspeaker under test.

If the signal used has a crest factor of 6 dB, and the loudspeaker is rated at 50 W continuous, the amplifier size required to run the test would be

$$\begin{aligned} 17\text{dBW} + 6\text{ dB} &= 23\text{ dBW} \\ 23\text{ dBW} &= 200\text{ watts continuous} \end{aligned}$$

Power specifications are of little use to system technicians. They must be converted to an equivalent rms voltage so that the system tech can measure the signal with a voltmeter. A simple conversion is to multiply the power rating by eight and take the square root to get the voltage. Bear in mind that if the power rating has been exaggerated, the voltage resulting from this conversion will be too.

Due to the high crest factors of audio program material, power amplifiers normally deliver far below their theoretical maximum sine wave power. This makes it possible (and necessary) to use an amplifier whose continuous power rating exceeds the continuous power rating of the loudspeaker if the intent is to produce the maximum L_P possible. Care is required on the part of the user to insure that the crest factor of the program material is not reduced excessively by dynamic range control devices (compressors and limiters) and then used to drive the amplifier to the point of clipping. Figs. 33-11 and 33-12 show the same waveform. The peak output voltage of each waveform is the same. A peak limiter was used to reduce the dynamic range of the second waveform, resulting in a 6 dB increase in applied rms voltage (and continuous power) to the loudspeaker (four times). This example shows how that amplifier power (and loudspeaker power dissipation) are highly dependent on the nature of the waveform, not just the amplifier rating. The amplifier selection and setting should ideally depend on the target sound level in the audience. There are no ramifications to operating a loudspeaker below its power rating, and in fact it is good design practice.

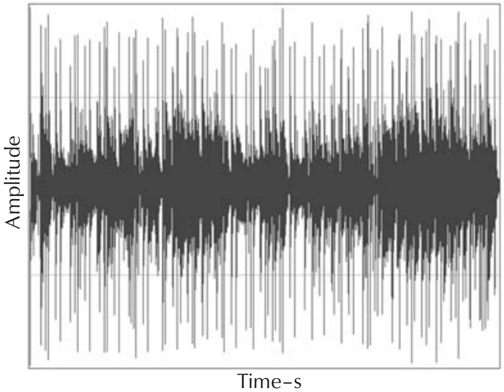


Figure 33-11. Output voltage of a complex waveform with large peaks.

One is reluctant to formalize a formula for determining the required amplifier size for several reasons, including:

1. The continuous output power of an amplifier is a function of the crest factor of the program material and can vary by 20 dB or more (a 100:1 ratio!).

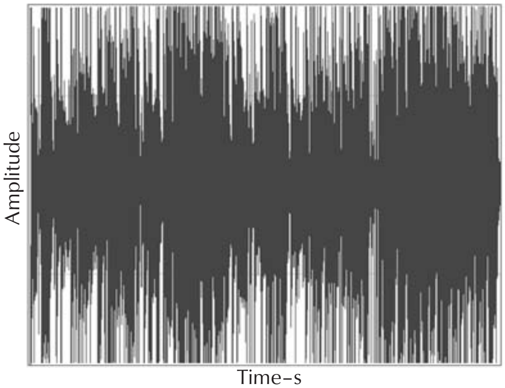


Figure 33-12. Output voltage of the waveform in Fig. 33-11 utilizing a 6 dB limiter and normalized to full scale.

2. Amplifier power ratings are based on signals that can bear little resemblance to audio program material.
3. Monitoring actual power delivered to the loudspeaker requires sophisticated equipment and a knowledgeable operator.
4. Standard loudspeaker power handling tests require that the loudspeaker be driven to the point where no permanent damage occurs. This is a bit ambiguous. The author utilizes a power handling test that drives the loudspeaker with increasing rms voltage until its response changes by 3 dB from the small signal (typically 3 Vrms) response. This rms voltage is used to determine the continuous rating of the loudspeaker, either in volts rms or power into a rated impedance.

Even so, a conservative approach is as follows:

1. Determine the loudspeaker's continuous power rating in watts (from the specification sheet). Determine the maximum rms voltage by taking the square root of eight times the power rating. Note that the voltage is necessary for level setting and verification.
2. Quadruple this rating for the required amplifier size. This will allow program peaks to exceed the continuous rating by 6 dB.
3. Be careful to not clip the amplifier, Fig. 33-13.

If the crest factor of the program material exceeds 6 dB, and the amplifier is operated without clipping, the loudspeaker will simply be operating further below its continuous rating, increasing its reliability and longevity. A careful operator could use a significantly larger amplifier, provided that a high crest factor is maintained and clipping is avoided, Fig. 33-14.

In essence, buy a big amplifier but use it carefully. Don't overdrive the loudspeaker or the audience! A sound level meter should always be used to check the L_P produced by the system, and this value should be within OSHA exposure guidelines.

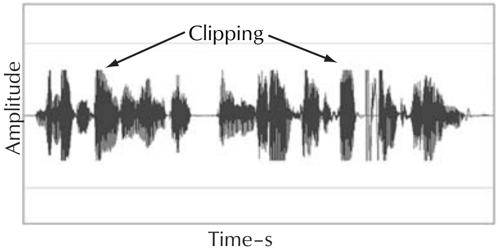


Figure 33-13. Waveform clipped due to insufficient head room.

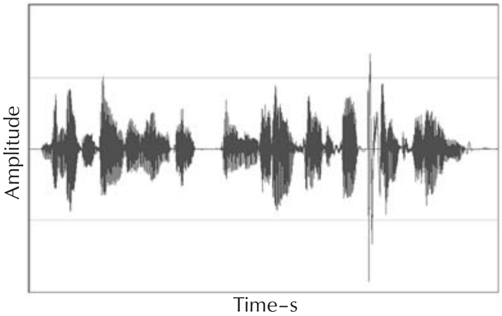


Figure 33-14. Increasing the amplifier size produces additional head room.

33.9 Conclusion

A properly calibrated sound system allows the operator to mix at or near meter zero on the mixer without danger of clipping any system component. Meter zero should also correlate with the maximum desired L_P in the audience. In effect, all components in the system are now functioning as one component, the only difference being that they are housed in separate chassis and interconnected with cables.

Reference

Milton Kaufman and Arthor Seidman, *Handbook for Electronics Engineering Technicians*, 1976 McGraw-Hill, Inc.
Don Davis and Eugene Patronis, *Sound System Engineering*, 2006 Focal Press, Burlington. MA.

Chapter 34

Sound System Design

by Chris Foreman

34.1 Introduction 1239
34.2 Sound Reinforcement System Models 1239
  34.2.1 The Four Questions (and a Fifth) 1239
  34.2.2 The Simplified (Outdoor) Sound Reinforcement System Model 1239
    34.2.2.1 Definitions 1240
    34.2.2.2 Attenuation with Increasing Distance 1240
    34.2.2.3 Acoustic Gain 1240
    34.2.2.4 Feedback and Potential Acoustic Gain (PAG) 1241
    34.2.2.5 Number of Open Microphones (NOM) 1241
    34.2.2.6 Feedback Stability Margin (FSM) 1241
    34.2.2.7 Noise 1241
    34.2.2.8 Head Room and Electrical Power Required (EPR) 1241
    34.2.2.9 Equivalent Acoustic Distance (EAD) 1242
    34.2.2.10 Needed Acoustic Gain (NAG) 1243
    34.2.2.11 Will the System Feed Back? 1243
    34.2.2.12 The Effect of Directional Microphones and Loudspeakers 1243
    34.2.2.13 How These Equations Answer the Four Questions 1244
34.2.2.13.1 Question 1: Is It Loud Enough? 1244
34.2.2.13.2 Question 2: Can Everybody Hear? 1244
34.2.2.13.3 Question 3: Can Everybody Understand? 1244
34.2.2.13.4 Question 4: Will It Feed Back? 1244
  34.2.3 Indoor Sound Reinforcement System Model 1245
    34.2.3.1 Echoes and Reverberation 1245
34.2.3.1.1 When an Echo Is a Problem 1245
34.2.3.1.2 Can a Reflection Be Useful? 1245
34.2.3.1.3 When Reverberation Is Useful 1245
34.2.3.1.4 When Reverberation Is a Problem 1246
    34.2.3.2 Reverberation and the Sound Reinforcement System 1246
34.2.3.2.1 Direct/Reverberant Ratio 1246
34.2.3.2.2 Loudspeaker Q 1246
34.2.3.2.3 Room Constant 1247
34.2.3.2.4 Critical Distance (Dc) 1247
    34.2.3.3 Attenuation of Sound Indoors 1248
    34.2.3.4 The Four Questions Again 1249
34.2.3.4.1 Question 1: Is It Loud Enough? 1249
34.2.3.4.2 Question 3: Can Everybody Understand? 1250
34.2.3.4.3 Alcons Modified for Audience Absorption and Talker/Listener Factors 1251
34.2.3.4.4 Alcons in Terms of RT_{60} and Direct, Reverberant, and Noise Levels 1251
34.2.3.4.5 Question 4: Will It Feed Back? 1252

34.2.3.4.6 The Effect of Directional Microphones and Loudspeakers 1253
34.2.3.5 Validity of the Model in a Geometrically Complex Room 1253
34.2.3.6 A Modification for Low RT60 Rooms 1254
34.3 Loudspeaker Systems for Sound Reinforcement 1255
34.3.1 Loudspeaker Components. 1255
34.3.1.1 Cone Loudspeakers 1255
34.3.1.2 Cone Loudspeaker Enclosures 1255
34.3.1.3 Compression Drivers and Horns 1256
34.3.1.4 Packaged Loudspeaker Systems 1257
34.3.1.5 Line Arrays 1257
34.3.1.6 Choosing Loudspeakers 1259
34.3.1.6.1 Power Handling 1259
34.3.1.6.2 Frequency Range and Response 1259
34.3.1.6.3 Sensitivity 1259
34.3.1.6.4 Coverage Pattern 1259
34.3.1.6.5 Evaluating Loudspeaker Sound Quality 1259
34.3.2 Loudspeaker Systems 1260
34.3.2.1 The Central Cluster 1260
34.3.2.2 Variations on the Central Cluster 1261
34.3.2.3 The Split Cluster 1261
34.3.2.4 Stereo Clusters 1261
34.3.2.5 Left-Center-Right Clusters 1261
34.3.2.6 Exploded Clusters 1262
34.3.2.7 Rear and Surround Clusters 1262
34.3.2.8 Designing a Central Cluster in a Simple Rectangular Room 1262
34.3.2.8.1 Evaluate the Room 1262
34.3.2.8.2 Choosing the Cluster Location 1263
34.3.2.8.3 Evaluating the Cluster Location 1263
34.3.2.8.4 Choosing the Loudspeakers 1263
34.3.2.8.5 Aiming the Loudspeakers 1263
34.3.2.8.6 Powering the Cluster 1263
34.3.2.8.7 Low-Frequency Loudspeakers and Feedback 1264
34.3.2.8.8 Direct Sound and Feedback 1264
34.3.2.8.9 Other Considerations 1264
34.3.2.9 Answering Question 2: “Can Everybody Hear” 1264
34.3.2.9.1 History of Manual/Graphical Cluster Design Tools 1264
34.3.2.9.2 Commercial Computer Software Design Tools 1265
34.3.2.9.3 Other Software Tools 1265
34.3.2.10 Designing the Complex Cluster 1265
34.3.2.10.1 The Fundamental Complexity 1266
34.3.2.10.2 Simplifying the Complex Cluster Design Process 1266
34.3.2.10.3 The Four Questions for the Complex Cluster 1267
34.3.2.10.4 Question 1: Is the System Loud Enough? 1267
34.3.2.10.5 Question 2: Can Everybody Hear? 1267
34.3.2.10.6 Question 3: Can Everybody Understand? 1267
34.3.2.10.7 Question 4: Will It Feed Back? 1267
34.3.2.11 Signal Alignment and Purposeful Misalignment in Cluster Design 1268
34.3.3 The Distributed Loudspeaker System. 1268
34.3.3.1 Distributed Ceiling Loudspeaker Systems 1269
34.3.3.2 Central Cluster Plus Distributed System 1269

34.3.3.3 Multicluster System 1270
34.3.3.4 Distributed Column System 1270
34.3.3.5 Pew-Back Distributed System 1270
34.3.3.6 Loudspeakers for Distributed Systems 1270
34.3.3.6.1 Ceiling Loudspeaker Systems 1270
34.3.3.6.2 Packaged Loudspeaker Systems in Distributed Systems 1271
34.3.3.7 Designing the Distributed Ceiling System 1271
34.3.3.7.1 Question 1: Is It Loud Enough? 1271
34.3.3.7.2 Question 3: Can Everybody Understand? 1272
34.3.3.7.3 Question 4: Will It Feed Back? 1272
34.3.3.7.4 Question 2: Can Everybody Hear? 1272
34.3.3.7.5 Equalizing the Distributed Ceiling System 1273
34.3.3.7.6 Distributed Systems in Rooms with Sloped Floors or Ceilings 1273
34.3.3.8 Designing the Distributed Cluster System 1274
34.3.3.8.1 Signal Delay in a Distributed Cluster System 1274
34.3.3.8.2 Distributed Clusters with No Delay 1274
34.3.3.8.3 Equalizing the Distributed Cluster System 1275
34.3.4 Crossover Networks and Biamplification 1275
34.3.4.1 Definitions 1275
34.3.4.2 Advantages of a Biamplified System 1275
34.3.4.3 When to Use a Loudspeaker System with Passive Crossover 1276
34.3.4.4 Signal Alignment in a Loudspeaker Crossover Network 1276
34.3.5 Protecting the Loudspeakers 1276
34.3.5.1 Loudspeaker Failure Modes 1276
34.3.5.1.1 Choosing Power Amplifiers to Prevent Excessive Average Power 1277
34.3.5.1.2 Loudspeaker Power Capacity Specifications 1277
34.3.5.2 Protecting against Excessive Low-Frequency Peak Power 1278
34.3.5.3 Loudspeaker Protection Devices 1278
34.3.5.3.1 Fuses 1278
34.3.5.3.2 Capacitors 1278
34.3.5.3.3 Limiters 1279
34.3.5.3.4 Other Protection Devices 1279
34.3.5.3.5 Protecting against Weather, Physical Damage and Pests 1280
34.3.5.3.6 Age-Related Loudspeaker Damage 1280
34.4 Electronic Components for Sound Reinforcement 1280
34.4.1 General Specifications fozr Sound Reinforcement 1280
34.4.2 Mixers and Mixing Consoles 1280
34.4.2.1 Mix Groups, Auxiliary Groups and Matrix Mixing 1281
34.4.3 Automatic Microphone Mixing 1282
34.4.3.1 Special Features 1284
34.4.3.2 Applications for Automatic Mixing 1284
34.4.3.3 Mix-Minus and Matrix Mixing 1285
34.4.3.4 Problems in Automatic Mixing 1285
34.4.4 Signal-Processing Components 1285
34.4.4.1 Compressors and Limiters 1286
34.4.4.1.1 Sound System Applications for a Compressor/Limiter 1286
34.4.4.1.2 Problems with Compressor/Limiters 1286
34.4.4.2 Equalizers 1286
34.4.4.2.1 Graphic Equalizers 1286
34.4.4.2.2 Boost-and-Cut Equalizers Versus Cut-Only Equalizers 1287

34.4.4.2.3 Constant Q versus Variable Q Equalizers 1287
  34.4.4.3 Parametric Equalizers 1287
  34.4.4.4 Digital and DSP Equalizers 1287
  34.4.4.5 High-Pass and Low-Pass Filters 1288
  34.4.4.6 Delay 1288
  34.4.4.7 Electronic Crossovers 1288
 34.4.5 Digital Signal Processing 1288
  34.4.5.1 Multifunction DSP 1288
  34.4.5.2 Power Amplifier DSP 1288
  34.4.5.3 Loudspeaker Processing DSP 1288
  34.4.5.4 Multifunction DSP in Sound Reinforcement System Design 1289
 34.4.6 Digital Audio Networking 1290
 34.4.7 Power Amplifiers 1290
 34.4.8 Pads and Transformers 1291
 34.4.9 70.7 Volt /100 Volt System Design 1291
  34.4.9.1 70 V Transformers 1291
  34.4.9.2 Designing a 70 V System 1292
34.5 System Installation and Commissioning 1292
 34.5.1 What Is Commissioning? 1292
  34.5.1.1 Step 1—Test All Components 1292
  34.5.1.2 Step 2—Adjust the Electronics 1292
  34.5.1.3 Step 3—System Adjustments and Equalization 1292
  34.5.1.4 Connectors and Cabling 1292
34.5.1.4.1 General Notes on Cable 1292
34.5.1.4.2 General Notes on Connectors 1293
34.5.1.4.3 XLR-Type Connection 1293
34.5.1.4.4 Phone Plugs 1293
34.5.1.4.5 RCA-Type Phono Plugs 1293
34.5.1.4.6 Barrier Block Connectors 1293
34.5.1.4.7 CAT 5 Connectors 1293
34.5.1.4.8 Cable and Connectors for Microphones and Other Low-Level Devices 1294
34.5.1.4.9 Microphone Snake Cables 1295
34.5.1.4.10 Cable and Connectors for Line-Level Devices 1295
34.5.1.4.11 Cable and Connectors for Loudspeakers 1295
  34.5.1.5 Cable and Connectors for Digital Audio Devices 1296
  34.5.1.6 Understanding Balanced and Unbalanced Lines 1296
  34.5.1.7 Impedance and Level Watching 1296
34.5.1.7.1 Terms: Source, Input, Output, Load 1296
34.5.1.7.2 Impedance Compatibility 1297
34.5.1.7.3 Passive Devices 1297
34.5.1.7.4 Passive Sources 1297
34.5.1.7.5 Active Devices 1298
34.5.1.7.6 Active Sources 1298
34.5.1.7.7 Impedance and Cable Length 1298
34.5.1.7.8 Signal-Level Compatibility 1298
34.5.1.7.9 The dBV Specification 1300
34.5.1.7.10 Impedance and Power Transfer 1300
  34.5.1.8 Digital Audio Level and Impedance Watching 1301
  34.5.1.9 Grounding and Shielding (also see Chapter 32) 1301
34.5.1.9.1 Grounding for Safety 1302

34.5.1.9.2 Improperly Wired ac Receptacles 1302
34.5.1.9.3 Two-Wire ac Receptacles 1302
34.5.1.9.4 Older Two-Wire Audio Equipment 1303
34.5.1.9.5 Grounding for Safety Outdoors 1303
34.5.1.9.6 Grounding to Reduce External Noise Pickup 1304
34.5.1.9.7 Definitions: Ground, Earth, Common 1304
34.5.1.9.8 Ground Loops in Unbalanced Systems 1304
34.5.1.9.9 Ground Loops in Balanced Systems and Using the Telescoping Shield Connection 1304
34.5.1.9.10 The Pin 1 Problem 1305
34.5.1.9.11 Transformers versus Active Balanced Inputs and Outputs 1305
34.5.1.9.12 Using Proper Shielding to Reduce Noise Pickup 1306
34.5.1.9.13 Reducing Noise Pickup from ac Lines 1306
34.5.1.9.14 More Tips on Reducing Noise Pickup 1306
  34.5.2 Testing and Adjusting 1307
    34.5.2.1 Signal Delay in Sound Reinforcement 1307
34.5.2.1.1 Signal Delay for Loudspeaker Clusters 1307
34.5.2.1.2 Signal Delay for an Under-Balcony Distributed System 1307
34.5.2.1.3 Signal Delay for a Loudspeaker or Loudspeaker Cluster 1307
    34.5.2.2 Equalization 1307
34.5.2.2.1 The Concept of Equalization 1307
34.5.2.2.2 What Equalization Can and Cannot Do 1308
34.5.2.2.3 System Design Criteria for Equalization 1308
34.5.2.2.4 System Equipment 1308
34.5.2.2.5 Test Equipment 1308
34.5.2.2.6 The Test Setup 1309
34.5.2.2.7 The Equalization Process 1309
34.5.2.2.8 The Desired System Response Curve 1310
34.5.2.2.9 Why High-Frequency Roll-Off Is Required in Equalization 1310
34.5.2.2.10 Use of High- and Low-Pass Filters 1310
34.5.2.2.11 Equalization Using TEF or Other Equipment That Measures Direct Sound Response 1311
34.5.2.2.12 Use of Narrow-Band Filters in Equalization 1311
34.5.2.2.13 Equalization Using Source-Independent Measurement (SIM) 1312
34.5.2.2.14 Automatic Equalization 1312
  34.5.3 Rigging the Cluster 1312
  34.5.4 System Documentation 1312
    34.5.4.1 Original Plans and As-Built Drawings 1314
    34.5.4.2 Equipment Lists and Equipment Owner's Manuals 1314
    34.5.4.3 Software Configurations and Backups 1314
    34.5.4.4 Approvals and Certifications 1314
    34.5.4.5 Additional Documentation 1314
  34.5.5 Troubleshooting a Sound System 1314
    34.5.5.1 Know the Block Diagram 1314
    34.5.5.2 Logical Troubleshooting 1314
    34.5.5.3 DSP Troubleshooting 1315
    34.5.5.4 Anticipating and Preparing for Problems 1317
34.6 Applications 1317
  34.6.1 Portable and Tour Sound Systems. 1317
    34.6.1.1 Packaging 1317
    34.6.1.2 Loudspeaker Systems 1317
    34.6.1.3 Portable System Rigging 1318

34.6.1.4 Cabling 1318
34.6.1.5 Electronics 1319
34.6.1.6 Multichannel Portable Systems 1319
34.6.1.7 Electrical and SPL Safety in Portable Systems 1319
34.6.1.8 Performance Criteria 1319
34.6.1.9 Stage Monitor Systems 1320
34.6.1.10 The Entertainment System as a Musical Instrument 1320
34.6.2 Systems for Religious Facilities 1320
34.6.2.1 User Interface 1321
34.6.2.2 Systems for Inexperienced Operators 1321
34.6.2.3 System Aesthetics 1321
34.6.2.4 Reverberant Room Problem 1322
34.6.3 Sports Stadiums and Other Outdoor Systems 1322
34.6.3.1 Excess Attenuation of High Frequencies in Air 1322
34.6.3.2 Using Distributed Systems Outdoors 1323
34.6.3.3 Echoes Outdoors, Artificial and Otherwise 1323
34.6.3.4 Dealing with Long Distances 1324
34.6.3.5 High Noise Outdoors (or Indoors) 1325
34.6.3.6 Dealing with Varying Noise Levels 1325
34.6.3.7 Thermal Layers of Air 1326
34.6.3.8 The Effect of Wind 1326
34.6.3.9 Dealing with Weather-Caused Deterioration 1327
34.6.3.10 Simulating the Indoor Environment Outdoors 1327
34.6.4 Artificial Ambience 1328
34.6.5 Conference and Boardroom Systems 1328
34.6.5.1 Local Sound Reinforcement in a Conference Room 1328
34.6.5.2 Recording a Conference 1329
34.6.5.3 Audio and Video Teleconferencing 1329
34.6.6 Sound Masking (Noise Masking), Speech Privacy Systems 1330
34.6.6.1 Criteria for the Environment 1330
34.6.6.2 Mechanical Noise Reduction 1330
34.6.6.3 Masking Sound 1330
34.6.6.4 Criteria for the Loudspeaker System 1330
34.6.6.5 Criteria for the Amplification System 1331
34.6.6.6 Criteria for the Equalizer 1331
34.6.6.7 A Second System 1331
34.6.6.8 Masking Plus Paging or Background Music 1331
34.6.6.9 Adjusting the Installed System 1332
34.6.6.10 Objective and Subjective Methods for Measuring System Effectiveness 1332
34.6.6.11 Conclusion 1333
34.7 Computers and Software for Sound Reinforcement Systems 1333
34.7.1 Digital Audio Devices. 1333
34.7.2 Audio Signal Transmission via Computer Networks 1333
34.7.3 System Control and Monitoring by Computer 1333
34.7.4 Computer Aided System Design. 1333
34.7.5 Computer Aided Measurement Systems. 1334
References 1334
Organizations and Web Sites 1335
Bibliography 1335

34.1 Introduction

There are a multitude of different types of sound systems with purposes as diverse as artificial ambience and voice paging, yet most share common design criteria. This chapter covers many of these design criteria using sound reinforcement systems as examples. Included are discussions of other types of systems such as foldback (stage) monitor systems and some types of playback only systems as well as some of the practical aspects of sound system design such as equipment choice and installation techniques.

Since the third edition of the *Handbook for Sound Engineers*, digital audio products, driven by DSP chips, have become mainstream choices for signal processing. Packaged loudspeaker systems and line arrays have replaced individual components for most loudspeaker cluster designs. And, most system designers use EASE or Bose Modeler or some other system design software in place of the traditional sound system analysis equations and graphical cluster design methods.

Yet, the original system analysis equations and graphical cluster design methods are the foundation for software programs like EASE and Modeler. An understanding of the original tools offers a modern designer a better understanding of the value, and the limitations, of these software tools.

For these reasons, and for historic completeness, this chapter retains its focus on the original sound system analysis equations, with revisions where appropriate. In addition, while the cluster design discussions use packaged loudspeaker systems or line arrays for most examples, component cluster designs are mentioned where they offer advantages. Finally, this chapter discusses how digital audio is changing system design as it replaces analog technology.

Most of the equations presented in this chapter can be used with either U.S. or metric distances. Just be consistent throughout. In a few cases, one or more constants change depending on the choice of units. The needed changes are noted near the equation. All of the examples use U.S. units.

34.2 Sound Reinforcement System Models

34.2.1 The Four Questions (and a Fifth)

The system design concepts presented in this chapter help to answer four simple questions:

- Question 1: Is it loud enough?
- Question 2: Can everybody hear?
- Question 3: Can everybody understand?
- Question 4: Will it feed back?

The sound reinforcement system models and equations in this chapter provide precise answers to these four questions and, in so doing, help the designer predict the success of the sound system in meeting its goals.

It is important to answer a fifth question, "Does it sound good?" The answer to this question may seem very subjective. However, good sound quality depends very much on favorable answers to the first four questions and these questions have objective answers. Also, good sound quality depends on other objective factors such as low distortion and smooth frequency response. Thus, while there are no equations to answer Question 5, it is possible to answer this question using objective criteria.

34.2.2 The Simplified (Outdoor) Sound Reinforcement System Model

Fig. 34-1 shows a simplified sound reinforcement system with a talker or sound source, a listener, a microphone, and a loudspeaker. (The somewhat awkward term, *talker*, replaces the term *speaker* to avoid confusion between the person talking and the loudspeaker system.) The talker could be replaced by a musician playing an instrument with no changes in the following discussions (although a very loud, amplified instrument or direct box connection would change things somewhat).

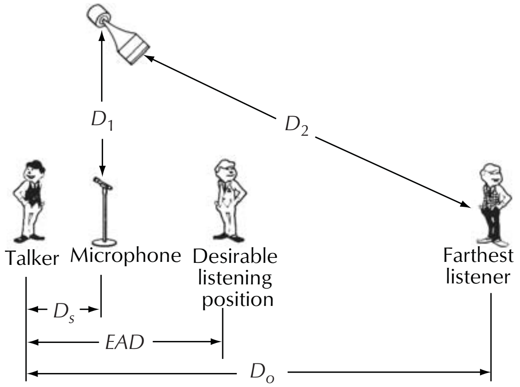


Figure 34-1. A simplified sound reinforcement system. Courtesy Bosch/Electro-Voice.

The primary simplification for this model is to ignore echoes and reverberation. This is a reasonable assumption for an outdoor system, away from any large buildings or other sources of echoes. As discussed later, this simplified description can be readily modified to include the effects of indoor reverberation.

34.2.2.1 Definitions

It is conventional to use the terms D_s, D_1, D_2, and D_0 when referring to the distances between the elements of this simple system.

- D_s.is the distance between the talker and the microphone.
- D_0 is the distance between the talker and the listener (again, the farthest listener when there are many listeners).
- D_1 is the distance between the loudspeaker and the microphone.
- D_2 is the distance between the loudspeaker and the listener (if there are many listeners, D_2 is usually considered to be the distance between the loudspeaker and the farthest listener).
- L_P—level of pressure—more commonly known as SPL (sound pressure level)

The terms D_s and D_0 start at the talker. The terms D_1 and D_2 are referenced to the loudspeaker. The first member of each pair measures to the microphone, D_s and D_1. The second member of each pair measures to the listener, D_0 and D_2. It's easy to remember them this way.

34.2.2.2 Attenuation with Increasing Distance

As the listener moves farther away from the loudspeaker, the sound pressure level L_P (or SPL) at the listener's ears will decrease. Neglecting the effect of echoes (outdoors, away from buildings this is a good approximation), the L_P will decrease exactly 6.02 dB every time the listener doubles the distance from the loudspeaker, Fig. 34-2). This effect is known as the "inverse-square law" and can be stated mathematically as

$$L_p{}' = L_p - 20\log\frac{D'}{D} \tag{34-1}$$

Example:

Let,

$L_P = 110$ dB,

D = 4 ft,

D′ = 200 ft.

$$L_p{}' = 110 - 20\log\frac{200}{4}$$

$$= 76.0 \text{ dB}$$

In Eq. 34-1, L_P is the sound pressure level (in dB SPL) the listener hears at distance D from the source (the loudspeaker in the simplified system). $L_P{}'$ is the new sound pressure level the listener would hear at distance D' from the source. If distance D' is smaller than distance D (the listener has moved closer to the source), then the term 20 log (D'/D) will be a negative number, and the new $L_P{}'$ will be greater than the original L_P. Note that because the equation uses a distance ratio, the distances may be measured in any convenient unit (feet, yards, meters) as long as both distances are measured in the same unit.

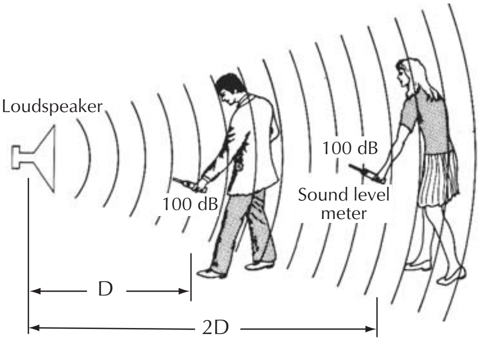


Figure 34-2. Inverse square law.

34.2.2.3 Acoustic Gain

Of all the reasons for a sound reinforcement system, the most important is implied by its name, sound reinforcement. That is, a sound reinforcement system reinforces a talker's voice so that the listener hears a louder sound with the system on than with the system off. The term *acoustic gain* describes the difference, in dB, between the sound pressure level, L_P, at the listener's ears with the system on and with the system off. In most cases, the listener means the farthest listener although the acoustic gain may be specified for any number of different listeners in a complex system. Acoustic gain may be described mathematically by a simple equation

$$\text{Acoustic gain} = L_{P_{on}} - L_{P_{off}} \tag{34-2}$$

Adequate acoustic gain is a primary design goal for a sound reinforcement system. Techniques for reaching this goal are described later. A simple technique for measuring the acoustic gain of a system is to place a sound level meter (SLM) at the position of the farthest listener and measure the L_P from the talker with the system off. Then turn the system on and measure the L_P

again. The difference between the two readings is the acoustic gain of the system. (Replace the talker with a pink-noise source, through a small loudspeaker, for a more consistent and accurate reading.)

34.2.2.4 Feedback and Potential Acoustic Gain (PAG)

The acoustic gain of this simple sound reinforcement system can be increased by turning up the volume control, but, at some point, this process will be interrupted by feedback (howling). Feedback is an undesirable oscillation of the entire sound reinforcement system that occurs when the sound from the loudspeaker feeds back to the microphone at a level high enough that the system begins to reinforce itself as well as it reinforces the talker.

Potential Acoustic Gain or *PAG* is the maximum acoustic gain that can be obtained from the system before feedback occurs. For this simplified system (neglecting reverberation and echoes), PAG can be stated mathematically as

$$PAG = 20\log\frac{D_0 D_1}{D_S D_2} \tag{34-3}$$

where,

D_s is the distance between the talker and the microphone,

D_1 is the distance between the loudspeaker and the microphone,

D_2 is the distance between the loudspeaker and the farthest listener,

D_0 is the distance between the talker and the farthest listener.

34.2.2.5 Number of Open Microphones (NOM)

This example system has only one microphone. Adding additional open (in-use) microphones increases the possibility of feedback and reduces the potential acoustic gain. The basic PAG equation Eq. 34-3 can be modified to include a number of open microphones (NOM) term as follows

$$PAG = 20\log\frac{D_0 D_1}{D_S D_2} - 10\log NOM\,. \tag{34-4}$$

34.2.2.6 Feedback Stability Margin (FSM)

Eq. 34-4 is theoretically correct but experience shows that a system operated at or very near its PAG will exhibit ringing and probably have an undesirable peaked frequency response. In addition, a sound system operated near its PAG will increase the effective room reverberation time in an indoor system. Thus, a 6 dB feedback stability margin (FSM) is normally subtracted from the calculated PAG. Systems operated 6 dB or more below their PAG are usually free of the problems of feedback or ringing. The final PAG equation for the simplified system, then, should include a FSM modifier as follows

$$PAG = 20\log\frac{D_0 D_1}{D_S D_2} - 10\log NOM - 6\text{ dB}\,. \tag{34-5}$$

Example:
Let,
$D_s = 2$ ft,
$D_0 = 128$ ft,
$D_1 = 45$ ft,
$D_2 = 90$ ft,
$NOM = 3$.

$$\begin{aligned} PAG &= 20\log\frac{128(45)}{2(90)} - 10\log 3 - 6\text{ dB} \\ &= 19.3\text{ dB} \end{aligned}$$

34.2.2.7 Noise

Unwanted noise (traffic, wind, audience noises, etc.) can interfere with the listener's ability to hear the talker. Ideally, the sound from the loudspeaker should be at least 25 dB above the noise level; that is, there should be a 25 dB SNR.

In some high-noise situations, a 25 dB SNR may not be achievable. Nevertheless, 25 dB is a common rule of thumb that will almost always insure that a listener can hear and understand the talker in an outdoor system.

34.2.2.8 Head Room and Electrical Power Required (EPR)

If ambient noise is 45 dB L_P (usually measured on the A scale of a sound level meter), and a 25 dB SNR is desired, then the desired L_P at the listener's ears is 70 dB. That 70 dB, however, is the average level at the listener's ears, and the peak L_P must be considered as well. The difference between peak and average level is referred to as the system *headroom*, Fig. 34-3. For a speech-only sound reinforcement system about 10 dB of head room is considered appropriate. Thus, the peak level in this example system would be 80 dB L_P for an average level of 70 dB L_P.

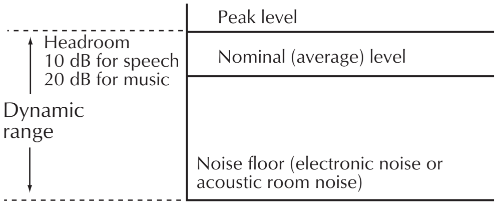


Figure 34-3. Dynamic range and head room.

In a high-noise system, a 10 dB head room factor may not be achievable. By using a limiter, the head room factor can be reduced to as low as 6 dB while maintaining reasonable voice intelligibility. For music-reinforcement systems, on the other hand, as much as 20 dB head room may be desirable to avoid clipping important musical peaks. For the simplified (outdoor) system, however, a 10 dB head room factor will be assumed.

How large an amplifier is needed to achieve the desired L_P? And, what information is needed about the loudspeaker? The answers are contained in the electrical power required (EPR) equation

$$EPR = 10^{\frac{L_P + H - L_S + 20\log\frac{D_2}{3.28}}{10}} \tag{34-6}$$

For metric distances, replace the constant 3.28 with the constant 1.00

where:

L_P is the average L_P required at distance D_2,
H is the head room in dB,
L_s is the sensitivity of the loudspeaker (1 W/1 m),
D_2 is the distance to the farthest listener.

Example:
Let
$L_P = 90$ dB,
$H = 10$ dB,
$L_s = 113$ dB (1 W/1 m),
$D_2 = 128$ ft.

$$\begin{aligned} EPR &= 10^{\frac{90 + 10 - 113 + 20\log\frac{128}{3.28}}{10}} \\ &= 76.3\text{ W} \end{aligned}$$

The term L_s is the rated sensitivity of the loudspeaker. This important specification represents the L_P that the loudspeaker will produce at one meter from its mouth with a one watt input power level and must be obtained from the manufacturer's specifications or from measurements performed in the field. Thus, this sensitivity is usually referred to as the "one-watt/one-meter" (1 W/1 m) sensitivity. In the past, some manufacturers used a one-watt/four-feet (1 W/4 ft) sensitivity rating.

The value of H (head room), of course, may be changed for a particular system, and a different D_2 could be used to find the EPR at some other distance.

34.2.2.9 Equivalent Acoustic Distance (EAD)

In the simplified system, if the talker were to stand relatively close to the listener, the talker could be heard and understood easily without the need for a sound system. One way of stating the goal of the sound system, then, is to say that it should create the illusion that the talker is close to the listener.

A simple experiment can determine just how close a talker needs to be to a listener for comfortable communication. Simply talk in a normal voice and walk backwards (away from the listener) until communication becomes difficult. Then walk toward the listener again until communication is comfortable. At this point, the equivalent acoustic distance (EAD) has been established. The idea is to use the sound system to create the illusion that the talker is this EAD away from the listener, Fig. 34-4).

In the simplified system, a 25 dB SNR is the goal. That means the L_P at the farthest listener's ears should reach 70 dB SPL for an assumed noise level of 45 dB SPL. Looking at the chart in Fig. 34-4 for a normal voice talker, and a 70 dB SPL (noise plus 25 dB SNR) level, the normal voice talker would have to be about 2 ft from the listener to achieve this desired L_P level. A raised voice talker would only have to be about 4 ft from the talker. Depending on the talker, one of these distances would be the required EAD. If the actual voice level from a talker (at some reference distance like one meter or four feet) is known, the EAD for the simplified system (outdoors) is as follows

$$EAD = D_S 10^{\frac{L_{Pt} - L_{Pd}}{20}} \tag{34-7}$$

where,

D_s is the reference distance from the talker,
L_{Pt} is the average L_P from the talker at distance D_s,
L_{Pd} is the desired L_P at the listener.

Example:
Let,
$D_s = 2$ ft,
$L_{Pt} = 71$ dB (at 2 ft),
$L_{Pd} = 65$ dB.

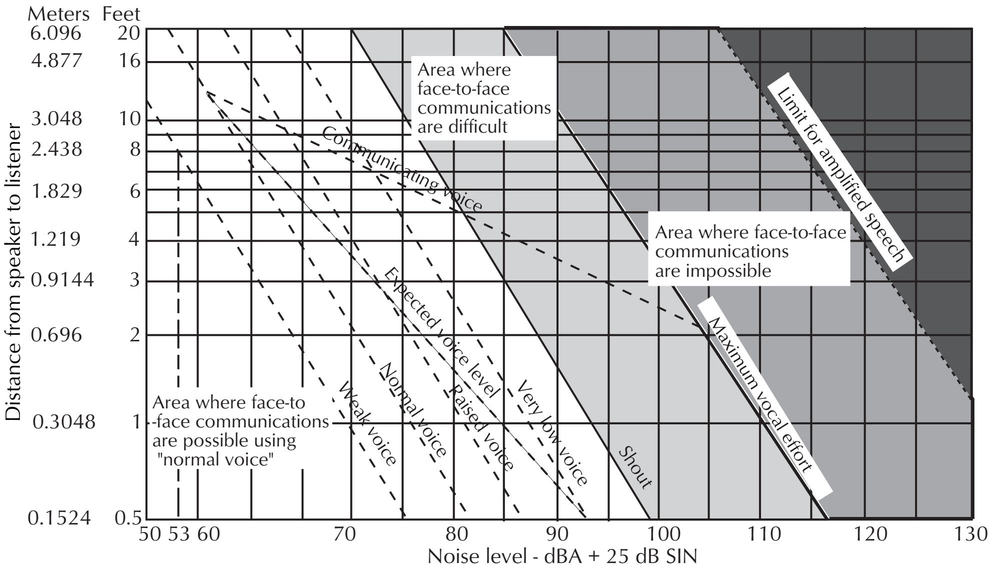


Figure 34-4. Nomograph for finding the EAD. Courtesy Syn-Aud-Con.

$$EAD = 2\left(10^{\frac{71-65}{20}}\right)$$

$$= 3.99 \text{ ft}$$

Here, the term D_s is the (reference) distance at which the L_P from the talker was measured. The term D_s (the microphone to talker distance discussed earlier) is used because it is a convenient number (normally about 1 m) and because it will make the next calculation (needed acoustic gain) easier. L_{Pt} is the sound pressure level from the talker at that reference distance D_s, and L_{Pd} is the desired sound pressure level at the listener (normally, this will be equal to the ambient noise plus a 25 dB SNR). EAD, then, is the equivalent acoustic distance number to be used in the next calculation, that of needed acoustic gain (NAG).

34.2.2.10 Needed Acoustic Gain (NAG)

The next question to be answered is, "How much acoustic gain is needed to achieve this desired L_P for a given talker's voice level?" This needed acoustic gain, or NAG, is the gain in decibels needed to produce the desired L_P at the listener's ears, L_{Pd}, given an *EAD* as calculated previously.

$$NAG = 20\log\frac{D_0}{EAD} \qquad (34\text{-}8)$$

where,

D_o is (as before) the distance from the talker to the farthest listener.

Example:
Let
$D_o = 128$ ft,
$EAD = 4$ ft.

$$NAG = 20\log\frac{128}{4}$$

$$= 30.1 \text{ dB}$$

34.2.2.11 Will the System Feed Back?

If the potential acoustic gain (PAG) from Eq. 34-5 is greater than or equal to the needed acoustic gain from Eq. 34-8, there is every reason to believe that the system will be stable and will not feed back. If, on the other hand, the potential acoustic gain is less than the needed acoustic gain, chances are good that the system won't work because turning up the volume control enough for the farthest listener to hear properly will always cause the system to be at or near feedback.

34.2.2.12 The Effect of Directional Microphones and Loudspeakers

The PAG and NAG equations assume an omnidirectional microphone and an omnidirectional loudspeaker. Some improvement in acoustic gain before feedback may be obtained by using a directional microphone (i.e., a cardioid pattern microphone) and/or by using a directional loudspeaker (a horn-type loudspeaker or line array). This only occurs if the D_1 is less than the critical distance D_c. (D_c is discussed in Section 34.2.3.2.4 and is only important in an indoor environment.)

A cardioid microphone could provide as much as 6 dB of additional gain before feedback if the rear of the microphone were pointed directly at the loudspeaker, as often happens with foldback stage monitor loudspeakers. The more typical case of a microphone at a podium and an overhead loudspeaker is a much less favorable arrangement since the side, not the rear, of the microphone will be pointed at the loudspeaker and D_1 is at or near D_c, providing 1 or 2 dB of additional gain before feedback at best. Because the microphone's cardioid pattern varies with frequency (it is nearly omnidirectional at low frequencies), even this 1 or 2 dB of feedback reduction may be optimistic. Thus, while a directional microphone may provide some additional gain before feedback, it's best to plan the system as if an omnidirectional microphone were to be used and take any additional gain provided by a cardioid microphone as a welcome bonus.

A directional loudspeaker may also provide some additional gain before feedback. For a horn-type loudspeaker pointed at the farthest listener, for example, the microphone will be off-axis of the horn, and the sound level at that off-axis angle may be –6 dB or more (down) compared to the on-axis level. This could theoretically provide an additional 6 dB gain before feedback. By using a highly directional horn, this 6 dB might be increased to 10 dB or even more. The fault with this theory is that there will almost always be a near-fill loudspeaker aimed at listeners near the microphone. This loudspeaker then becomes the limiting factor in the feedback loop. Even when the nearest listeners are far enough away from the microphone that the near-fill loudspeaker can be aimed well away from the microphone, the system woofer remains a potential feedback problem-causer. A highly directional low-frequency horn is physically very large and, thus, is seldom used in a cluster. A smaller low-frequency horn or vented, box-type, low-frequency component will be almost omnidirectional at low frequencies. Thus, the use of directional horns usually cannot improve gain before feedback because the feedback problems simply shift to below the crossover frequency.

A line array may also provide some additional gain before feedback. A properly designed and installed line array has a narrow vertical dispersion over a wide frequency range. This can keep the sound aimed at the audience and away from the system microphone. At some low frequency, however, the line array's ability to control its vertical dispersion is degraded and feedback may become an issue below this frequency. Also, it is not always possible to position the line array in such a way to avoid all feedback issues. For example, column-style line arrays are commonly placed to the left and right of the system microphone where their vertical directional control offers little advantage in controlling feedback.

34.2.2.13 How These Equations Answer the Four Questions

There are many other things to consider for even a simplified system. Outdoors, for example, there are the effects of wind, humidity, and temperature layers. (See Chapter 2 and Section 34.6.3.7.) However, the answers to the four questions supply a great deal of information about whether or not a system will actually reinforce sound in a satisfactory manner.

34.2.2.13.1 Question 1: Is It Loud Enough?

Eq. 34-6 helps answer Question 1, "Is it loud enough?." This equation doesn't take gain or feedback into account, however. Those concerns are covered in Question 4.

34.2.2.13.2 Question 2: Can Everybody Hear?

This question "Can everybody hear?" involves the coverage patterns of the loudspeakers and the way they are aimed into the audience. This topic is covered in detail in Section 34.3.2 and in Chapters 18 and 35.

34.2.2.13.3 Question 3: Can Everybody Understand?

For the simplified (outdoor) system, the answer to this question is yes if the system is loud enough (Question 1) and if it avoids problems like very poor frequency response and excessive distortion. Indoors, this question also involves the effects of reverberation.

34.2.2.13.4 Question 4: Will It Feed Back?

The answer to the question "Will it feed back?" is no (the desired answer) if PAG is equal to or greater than NAG, Eqs. 34-5 and 34-8.

Again, there's a lot more to sound system design than answering just four questions, but these particular four questions are important enough that, in the next section, they form the basis for an expansion of the simplified system to cover the effects of reverberation on indoor systems.

34.2.3 Indoor Sound Reinforcement System Model

So far, the discussion of sound reinforcement has been simplified by neglecting the effects of echoes and reverberation. Now, however, it's time to modify the mathematical model of the sound reinforcement system to include these effects. Doing this, of course, creates a much more useful model, one that can be used successfully both indoors and out.

The equations presented in this section used to determine indoor attenuation, critical distance, potential and needed acoustic gain, and electrical power required are derived from concepts presented by Hopkins and Stryker. The equation for *Alcons* was derived by Peutz and Klein. The acoustic gain (potential, and needed) equations were developed by Don Davis of Syn-Aud-Con. The critical distance (D_c) equation was developed by Don Davis and Mel Sprinkle. The equations have been manipulated and modified by a number of writers to make them more useful to sound system designers. The most notable of these writers are Don and Carolyn Davis.

The equations presented here are basically the same as those used in *Sound System Engineering* by Don and Carolyn Davis.[4] However, as presented here, the equations have been algebraically manipulated by this author to make them somewhat easier to explain and to make them more adaptable to computer analysis. As in the simplified model, the equations in this section help the designer answer the four questions.

34.2.3.1 Echoes and Reverberation

Echoes and reverberation are both reflections of sound. A reflection is called an echo if the time between the original sound and the reflection is long enough that both sounds can be heard distinctly (about 70 ms or greater). If a room has lots of reflections and they are closely spaced in time so that distinct echoes are not audible, this large number of reflections is known as reverberation. A much more detailed discussion of echoes, reverberation, and general room acoustics can be found in Chapters 1, 3, 5, and 7.

34.2.3.1.1 When an Echo Is a Problem

Some rooms have one or more distinct echoes but very little reverberation. A conference room with carpeting, draperies, padded seating, and acoustical ceiling tile, for example, may have little or no reverberation. That same room, however, may have a hard rear wall that produces a single slap-back echo (so called because it slaps back at talkers every time they try to speak from a location in the front of the room). Other, larger, rooms may have multiple distinct echoes. Superdome-sized rooms are an obvious example. In most cases, problem echoes must be dealt with by acoustic treatment. In some cases, in fact, a sound system will only aggravate an echo problem.

34.2.3.1.2 Can a Reflection Be Useful?

Reflections add to the level perceived by the listener but this additional level may or may not be useful. Mid- to high-level late reflections, which arrive at the listener's ear more than 50 ms after the direct sound, can muddy the sound or may even be perceived as echoes. Mid- to high-level early reflections, which arrive at the listener's ear less than 20 ms after the direct sound, cause comb filtering which degrades the system frequency response. Reflections between 20 ms and 50 ms, however, can add to the level in a way that is beneficial to intelligibility and pleasing to the sound quality.

Thus, audible reflections can be useful, questionable, or undesirable, depending primarily on the difference in the arrival time at the listener's ears (and somewhat on the difference in level between the direct and reflected sound).

Note that the sound from multiple loudspeakers can arrive at a listener's ears at multiple different times. This effect mimics early reflections and may or may not be useful as discussed above.

34.2.3.1.3 When Reverberation Is Useful

Some reverberation is often desirable, especially for a musical performance. The reverberation of a large cathedral, for example, enhances the organ and choir sound. Some musical compositions, like those written for a pipe organ, are actually intended for a large, reverberant room.

A small amount of reverberation can also enhance a speech reinforcement system. Reverberation can fill out a vocal sound to make it more natural. Those reflections in a reverberant field that reach the listener's ears a short time (but not too short a time) after the source can also, as explained previously, improve the ability to understand speech by effectively making it louder.

34.2.3.1.4 When Reverberation Is a Problem

While some reverberation is useful, too much reverberation causes muddy or boomy sound quality in a musical performance and makes speech very difficult to understand. In this case, the reverberation is a problem similar to that of too much ambient noise except that the reverberation gets louder as the signal gets louder so that a reverberation problem cannot be solved by simply turning up the volume control.

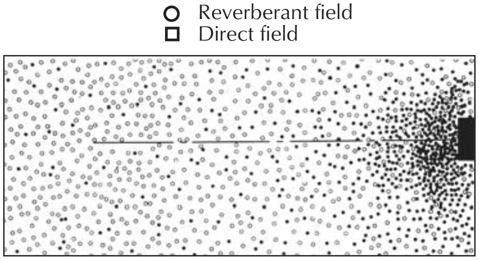
34.2.3.2 Reverberation and the Sound Reinforcement System

The indoor sound system model includes two important assumptions:

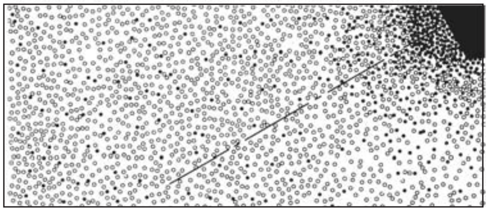
1. The room has no distinct echoes.
2. The room has a well-developed and statistically random reverberant field, Fig. 34-5.

The first assumption limits the model to rooms with few or no distinct echoes. This is an acceptable limitation since a room with distinct echoes needs acoustic treatment before a sound system can be applied. The second assumption is acceptable for the purposes of this section, although it should be understood that differences in reflecting and absorbing surfaces in most rooms prevent true randomness.

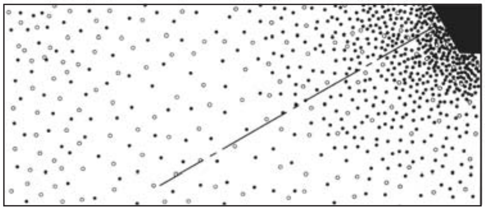
The reason the reverberant field must be considered is that it will help the system maintain a more consistent L_P from seat to seat, even though it will hinder, to some extent, the attempt to provide intelligible sound to every seat.



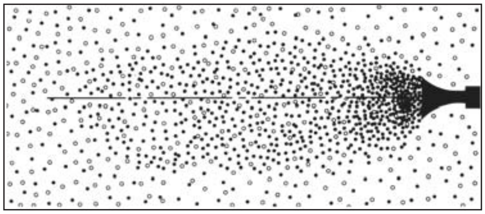
A. Nondirectional loudspeakers.



B. Loudspeaker in "live" rooms.



C. Loudspeaker in "dead" rooms.



D. Directional loudspeakers.

Figure 34-5. Direct and reverberant fields in a room. Courtesy JBL Professional.

34.2.3.2.1 Direct/Reverberant Ratio

Direct/reverberant ratio is a ratio of the direct sound at some point in a room to the reverberant sound, which is assumed to be the same everywhere in the room. A high direct/reverberant ratio means good speech intelligibility (if all other factors are favorable).

34.2.3.2.2 Loudspeaker Q

Q is a measure of the directional properties of a loudspeaker (also see Chapters 17 and 18). An omnidirectional loudspeaker has a Q of 1. A loudspeaker radiating into a hemisphere has a Q of 2. A loudspeaker radiating into half a hemisphere has a Q of 4 and so on, as shown in Fig. 34-6. A related term, DI (directivity index), is simply ten times the log (base 10) of Q. That is

$$DI = 10\log Q \qquad (34\text{-}9)$$

A loudspeaker with a high Q will have a narrow coverage pattern and will, therefore, concentrate its sound energy on fewer seats than a low-Q loudspeaker. Thus, the high-Q loudspeaker can provide higher levels of direct sound and, likewise, higher direct/reverberant ratios. This leads to better intelligibility, at least in the single loudspeaker case.

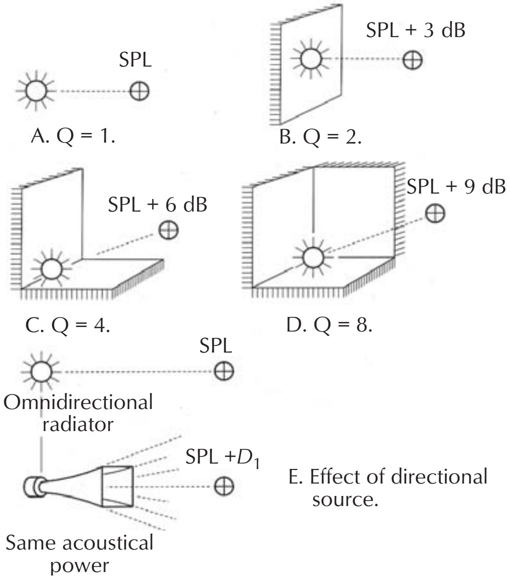


Figure 34-6. Directivity, angular coverage, directivity index (DI) and directivity factor (Q). Courtesy JBL Professional.

Q compares the on-axis sound intensity of a single loudspeaker to what that intensity would be if the loudspeaker were omnidirectional. Note that sound intensity is proportional to the sound pressure L_P squared. Because *Q* is only defined for a single loudspeaker, any mention of the off-axis *Q* or the *Q* of a cluster is technically inaccurate. A value for off-axis *Q* is useful, however, in calculating *Alcons* (articulation loss of consonents) for a listener not directly on-axis of a loudspeaker. A value for the *Q* of a cluster could be useful to help determine *Alcons* for a listener seated on-axis of a single horn in a multihorn system. Thus, the concept of *Q* is often extended to include these and other ideas.

The off-axis Q of a horn, for example, can be determined from an examination of the on-axis DI and the difference in sound pressure level on axis versus off-axis at the angle of interest. Subtract this difference from the on-axis DI and convert back to *Q*. For example, for a 90° (horizontal) horn, at 45° off-axis horizontally, the SPL is –6 dB from its on-axis value. Thus, if the on-axis DI is 12, the off-axis DI will be 6. In general, if DI is known

$$Q = 10^{\frac{DI}{10}} \tag{34-10}$$

Thus, the off-axis *Q* in this example is approximately 2.

The concept of the *Q* of a cluster is more difficult. In theory, it would be possible to calculate the *Q* of a cluster for a listener seated at any point in the room by comparing the direct L_P at that position with the overall acoustic power output of the entire cluster. In practice, this is a complex calculation since it requires a detailed knowledge of the efficiency of each loudspeaker and the electrical input to each driver as well as the directional characteristics and efficiency of each horn and any alterations that may be made in the horn's directional characteristics by the baffling effects of the cluster. One way to deal with these problems is discussed in Section 34.3.2.10.2, where the calculation of D_c modifiers is discussed.

34.2.3.2.3 Room Constant

Room constant, *R* (or *Sa*), is a measure of the relative liveness of a room (a live room has a well-developed, very audible reverberant field). A low room constant, or low *Sa*, means a very live room. A high *R* or *Sa* means a dead room. The *R* or *Sa* value depends on the size of the room, so a specific value of *R* or *Sa* is not enough to judge the reverberation characteristics of a room. Mathematically, room constant may be calculated as

$$R = \frac{S\bar{a}}{1-\bar{a}} \tag{34-11}$$

where,

S is the total surface area of the room,

$\bar{a}$ is the average absorption constant.

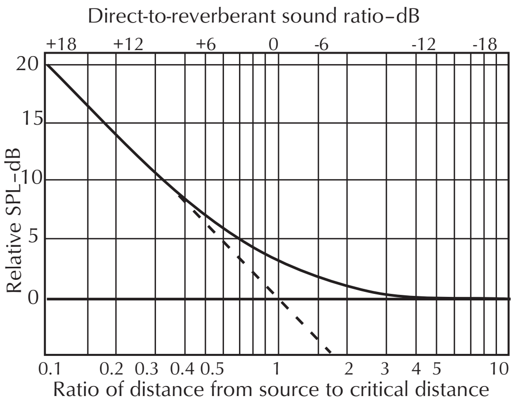
One version of the equation for critical distance uses the room constant in place of the *Sa* term. While room constant was commonly used to specify a room in the past (in the D_c equation), it has fallen into disuse and is usually replaced in most equations by the *Sa* term.

34.2.3.2.4 Critical Distance (D_c)

D_c (critical distance) is the distance from a source at which the direct sound is exactly the same L_p as the reverberant field, Fig. 34-7. Critical distance is important in a number of concepts including intelligibility.

A good estimate of the critical distance for a loudspeaker in a given room can be made by playing a pink-noise source through the loudspeaker and walking away from it holding a sound level meter. At some distance, the L_P will cease to change. Now walk back toward the source until the L_P increases exactly 3 dB. That distance will be the critical distance. (Since the direct sound and reverberant sound are equal here, the total is 3 dB above the reverberant sound alone.) Crit-

if



Calculated from 10 log (1+1/X²) where X is the ratio of distance from source to critical distance

Figure 34-7. Critical distance. Courtesy JBL Professional.

ical distance depends on the Q of the source and the absorption in the room; thus, it can vary with frequency, and this test shows only a broadband approximation of critical distance.

For a given loudspeaker in a given room, critical distance can be found from

$$\begin{aligned} D_c &= \sqrt{\frac{QS\bar{a}}{16\pi N}} \\ &= 0.141\sqrt{\frac{QS\bar{a}}{N}} \end{aligned} \tag{34-12}$$

where,

Q is the Q of the source,

S is the total surface area of the room,

$\bar{a}$ is the average absorption coefficient for the surfaces in the room,

N is the total number of loudspeakers producing the same acoustic power as the loudspeaker pointed at the farthest listener.

Example:
Let
$Q = 5$,
$S = 28{,}000\ \text{ft}^2$,
$\bar{a} = 0.35$,
$N = 1$.

$$\begin{aligned} D_c &= \sqrt{\frac{5(28,000)0.35}{16\pi 1}} \\ &= 31.2\ \text{ft} \end{aligned}$$

For a more detailed discussion of the concept of N, see Section 34.3.2.10.

34.2.3.3 Attenuation of Sound Indoors

The first part of the indoor sound reinforcement system model will tell us what happens to a sound at increasing distances from the source in an indoor environment. The inverse-square law, Eq. 34-1, is still correct indoors, but only for the direct sound. The reverberant sound level is assumed to be the same everywhere—that is, the reverberant sound level does not change with distance from the source. Thus, the total sound level, at any distance from a source, is the sum of the direct sound, which has been attenuated by inverse-square law, and the reverberant sound, which does not change with distance

$$L_p{}' = L_P - \left(20\log\frac{D'}{D} + 10\log\frac{g(D')}{g(D)}\right) \tag{34-13}$$

where,

D is the original distance from the source,

D' is the new distance from the source,

L_P is the original L_P at D,

$L_P{}'$ is the new L_P at the distance D',

$g(x)$ is found from the equation

$$g(x) = D_c{}^2 + x^2 \tag{34-14}$$

where,

x is any distance.

Note that the equation for indoor attenuation is exactly the same as Eq. 34-1 for the simplified system (outdoor) attenuation (inverse-square law) except for the final term, which can be interpreted as a contribution from the indoor reverberant field.

Example:
Let
$L_P = 90$ dB,
$D = 4$ ft,
$D' = 125$ ft,
$D_c = 31.2$ ft.

$$\begin{aligned} L_p{}' &= 90 - 20\log\frac{125}{4} + 10\log\frac{g(125)}{g(4)} \\ &= 72.3\ \text{dB}. \end{aligned}$$

To compare to outdoor attenuation, Eq. 34-1, simply ignore the term

$$10\log\frac{g(D')}{g(D)}$$

Indoor attenuation can also be found from another equation

$$L_p' = L_P - \Delta dB \tag{34-15}$$

where,

$$\Delta dB = \Delta D' - \Delta D$$

and

$$\Delta x = -10\log\left(\frac{Q}{4\pi x^2} + \frac{4}{S\bar{a}}\right)$$

where,
x is any distance.

Example:
Let,
$L_P = 90$ dB,
$D = 4$ ft,
$D' = 125$ ft,
$Q = 5$,
$S = 28{,}000$,
$\bar{a} = 0.35$.

$$\Delta D = -10\log\left(\frac{5}{4\pi 4^2} + \frac{4}{28{,}000\ (0.35)}\right)$$
$$= 16\ \text{dB}$$

$$\Delta D' = -10\log\left(\frac{5}{4\pi 125^2} + \frac{4}{28{,}000\ (0.35)}\right)$$
$$= 33.6\ \text{dB}$$

then,

$$\Delta\ \text{dB} = 33.6\ \text{dB} - 16\ \text{dB}$$
$$= 17.6\ \text{dB}$$

and

$$L_P' = 90 - 17.6$$
$$= 72.4\ \text{dB}.$$

Note: Except for round-off errors, this answer and the answer to the example for Eq. 34-14 are the same.

Eq. 34-15 is more common in the literature but Eq. 34-14 may be easier to understand and to use in a computer program. The two equations are mathematically the same and will produce the same answers given the same data.

34.2.3.4 The Four Questions Again

Question 2: "Can everybody hear?" is discussed in Section 34.3.2. Questions 1, 3, and 4, however, can be answered with the information obtained so far.

34.2.3.4.1 Question 1: Is It Loud Enough?

In the simplified (outdoor) system, the answer to this question depended on the required L_P at the farthest listener (at D_2), the required head room in decibels and the sensitivity of the loudspeaker. The answer was given in terms of the required electrical power to be supplied to the loudspeaker (the *EPR*).

In the indoor system, reverberation in the room affects the analysis. Yet, although the room adds complexity to the answer to Question 1, it makes things a little easier in the actual design of an indoor sound reinforcement system, because, after the critical distance is passed, the sound can only attenuate another 3 dB. Thus, for distances beyond D_c, no more power is needed to maintain the same L_P. Unfortunately, intelligibility suffers at distances well into the reverberant field. But that is the topic of Question 3. For now, here is the equation for indoor electrical power required

$$EPR = 10^{\frac{L_p + H - L_S + 20\log\frac{D_2}{3.28} - 10\log\frac{g(D_2)}{g(3.28)}}{10}} \tag{34-16}$$

where,
L_P is the average L_P required at distance D_2,
H is the head room in dB,
L_s is the sensitivity of the loudspeaker (1 W/1 m),
D_2 is the distance to the farthest listener,
For metric distances, replace the constant 3.28 with the constant 1.00.

Example:
Let
$L_P = 90$ dB,
$H = 10$ dB,
$L_s = 113$ dB (1 W/1 m),
$D_c = 31.2$ ft,
$D_2 = 128$ ft.

$$EPR = 10^{\frac{90 + 10 - 113 + 20\log\frac{128}{3.28} - 10\log\frac{17{,}357}{989.4}}{10}}$$
$$= 4.3\ \text{W}$$

To compare with the simplified system (outdoor) EPR equation (Eq. 34-6), simply ignore the term

$$10\log\frac{g(D_2)}{g(3.28)}.$$

The value L_P in the first line of the equation is the desired L_P at D_2. The value H is a desired value for head room, usually assumed to be 10 dB. That 10 dB, of course, may be changed for a particular system. D_c is the critical distance given in Eq. 34-12. It is instructive to note that this equation is the same as the simplified system (outdoor) EPR equation (Eq. 34-6) except for the term

$$10\log\frac{g(D_2)}{g(3.28)}$$

which can be interpreted as a contribution from the reverberant field.

EPR can also be found from an equation that is more common in the literature

$$EPR = 10^{\frac{L_p + H - L_S + \Delta D_2 - \Delta 3.28 - L_S}{10}} \qquad (34\text{-}17)$$

where,

$$\Delta x = -10\log\left(\frac{Q}{4\pi x^2} + \frac{4}{S\bar{a}}\right)$$

where,

x is any distance,

other terms are as before,

For metric distances, replace the constant 3.28 with the constant 1.00.

Example:

Let

$L_P = 90$ dB,

$H = 10$,

$D_2 = 128$ ft,

$L_s = 113$ dB/1 W/1 m,

$Q = 5$,

$S = 28{,}000$,

$\bar{a} = 0.35$.

then

$$\Delta D_2 = 33.6$$

and

$$\Delta 3.28 = 14.3$$

and

$$EPR = 10^{\frac{90 + 10 + 33.6 - 14.3 - 113}{10}}$$
$$= 4.33\ \text{W}$$

Eq. 34-17 may be more familiar to some readers. However, it and Eq. 34-16 are mathematically equivalent and will produce the same answers from the same data. Eq. 34-16, however, may be easier to understand and to use in a computer program.

34.2.3.4.2 Question 3: Can Everybody Understand?

In the simplified (outdoor) system, Question 3 was answered by considering the required SNR and making certain that the sound system output was sufficiently above the ambient noise level to provide intelligible sound (speech). Indoor intelligibility also depends on the reverberation time and the direct/reverberant ratio, and an unfavorable reverberation time or direct/reverberant ratio cannot be made better by merely increasing the SPL from the loudspeakers, since that will also increase the reverberant field level!

If the reverberation time or direct/reverberant ratio is unfavorable, one or more of the following may help:

1. Decrease the reverberant field by adding absorption to the room (usually a costly process).
2. Move the listener closer to the loudspeaker (in a reverberant church with the pews half filled, people sitting near the loudspeakers will hear and understand better than those farther away from the loudspeakers).
3. Move the loudspeakers closer to the listeners (adding additional loudspeakers, as described later, is a common way to improve direct-to-reverberant ratio but this is not a panacea since the new loudspeakers will add to the reverberant level as well as the direct level).
4. Use a loudspeaker with higher Q (this is ideal provided the required Q doesn't mean that you have a very narrow coverage pattern that cannot cover all the listeners).

How is a direct/reverberant ratio determined? What is a favorable reverberation time or direct/reverberant ratio? It is possible to determine this ratio directly, but it is more common to use the articulation loss of consonants concept. If the *Alcons* is 10% or less and the *SNR* is favorable (+15 dB or greater), there is every reason to believe the answer to Question 3 will be yes.

Note that some texts suggest an *Alcons* of 15% or less is acceptable. Also note that this *Alcons* equation is most accurate in rooms with reverberation times of 1.6 s or longer and is not particularly accurate or useful in acoustically dead rooms. In addition, this form of the *Alcons* equation assumes at least a 25 dB SNR. See Chapter 36 for a thorough discussion of speech intelligibility including additional speech intelligibility measurement standards.

$$Alcons = \frac{656 D_2{}^2 RT_{60}{}^2 N}{QV} \tag{34-18}$$

where,

D_2 is the distance between the loudspeaker and the farthest listener,

RT_{60} is the room reverberation time,

N is a number that attempts to compensate for the fact that there are most likely several loudspeakers in the cluster and only one will be pointed at the farthest listener (but all add to the reverberant field). Note that some texts use $N + 1$ in place of N,

Q is the Q of the loudspeaker,

V is the volume of the room,

For metric distances, replace the constant 656 with the constant 200,

D^2 is in m^2 for the SI system.

Example:

Let,

$D_2 = 125$ ft,

$RT_{60} = 2.5$ s,

$V = 500{,}000$ ft^3,

$Q = 10$,

$N = 1$.

$$Alcons = \frac{656(125^2)(2.5^2)(1)}{10(500,000)}$$
$$= 12.8\% \text{ (less than exceptable)}$$

One way to interpret N is the total number of loudspeakers producing the same acoustic power as the loudspeaker pointed at the farthest listener. For a more detailed discussion of the concept of N, see Section 34.3.2.10.

34.2.3.4.3 Alcons Modified for Audience Absorption and Talker/Listener Factors

The following equation for *Alcons* includes modifiers for the absorption of an audience and for a talker with poor articulation or listeners with less than normal hearing acuity. Use this form of the equation when listener or talker difficulties are expected or when the audience area will be significantly more absorptive than the rest of the room such as a large cathedral with marble walls and ceiling but with pew cushions in the seating area and carpeting on the floor.

$$Alcons = \frac{656 D_2{}^2 RT_{60}{}^2 N}{QV} + k \tag{34-19}$$

where,

$$m = \frac{1-a}{1-ac},$$

a is the average absorption coefficient for the room,

ac is the absorption coefficient in the area covered by the loudspeaker (the audience area),

k is a correction factor for a talker with poor articulation or listeners with less than normal hearing acuity and is typically in the range of 1–3%,

other factors are as before.

For metric distances, replace the constant 656 with the constant 200,

D^2 is in m^2 and V is in m^3 in the metric system.

34.2.3.4.4 Alcons in Terms of RT_{60} and Direct, Reverberant, and Noise Levels

The following equation presents *Alcons* in terms of the four factors most critical to intelligibility, which are direct sound level, reverberant sound level, reverberation time, and noise level. Use this form of the equation if the four quantities are known or can be reliably estimated.

$$Alcons = 100(10^{-2[A + BC - ABC]} + 0.015) \tag{34-20}$$

where,

$$A = -0.32\log\left(\frac{E_R + E_N}{10E_D + E_R + E_N}\right),$$

$$B = -0.32\log\left(\frac{E_N}{10E_R + E_N}\right),$$

$$C = -0.5\log\left(\frac{T_{60}}{12}\right),$$

$$E_R = 10^{\frac{L_R}{10}},$$

$$E_D = 10^{\frac{L_D}{10}},$$

$$E_N = 10^{\frac{L_N}{10}},$$

L_R is the reverberant level in the 2 kHz octave band,
L_D is the direct sound level in the 2 kHz octave band,
L_N is the ambient noise level in the 2 kHz octave band.

34.2.3.4.5 Question 4: Will It Feed Back?

The PAG and NAG concepts work indoors, too, but are modified by the room.

$$PAG = 20\log\frac{D_0D_1}{D_sD_2} - 10\log NOM - 6\text{ dB} - 10\log\frac{g(D_0)g(D_1)}{g(D_s)g(D_2)} \tag{34-21}$$

where,

$$g(x) = D_c^{\,2} + x^2,$$

other terms are as before.

Example:
Let,
$D_s = 2$ ft,
$D_o = 128$ ft,
$D_1 = 45$ ft,
$D_2 = 90$ ft,
$D_c = 31.2$ ft,
$NOM = 3$.

$$\begin{aligned} PAG &= 20\log\frac{128(45)}{2(90)} - 10\log 3 - 6\text{ dB} - 10\log\frac{17{,}357\ (2998)}{977(9073)} \\ &= 11.7\text{ dB}. \end{aligned}$$

To compare to the simplified (outdoor) system PAG Eq. 34-5, simply ignore the term

$$-10\log\frac{g(D_o)g(D_1)}{g(D_S)g(D_2)}$$

$$NAG = 20\log\frac{D_o}{EAD} - 10\log\frac{g(D_o)}{g(EAD)} \tag{34-22}$$

where,

$$g(x) = D_c^{\,2} + x^2,$$

other terms are as before.

Example:
Let
$D_o = 128$ ft,
$EAD = 4$ ft,
$D_c = 31.2$ ft.

$$\begin{aligned} NAG &= 20\log\frac{128}{4} - 10\log\frac{g(17{,}357)}{g(989)} \\ &= 17.7\text{ dB} \end{aligned}$$

To compare to Eq. 34-8, simply ignore the term

$$-10\log\frac{g(D_o)}{g(EAD)}$$

Alternate forms of the PAG and NAG equations, which are more common in the literature, follow:

$$PAG = \Delta D_o + \Delta D_1 - \Delta D_s - \Delta D_2 - 10\log NOM - 6\text{ dB} \tag{34-23}$$

where,

$$\Delta x = -10\log\left(\frac{Q}{4\pi x^2} + \frac{4}{S\bar{a}}\right)$$

where,
x is any distance,
other terms are as before.

Example:
Let
$D_o = 128$ ft,
$D_1 = 45$ ft,
$D_s = 2$ ft,
$D_2 = 90$ ft,
$Q = 5$,
$S = 28{,}000$ ft^2,
$a = 0.35$,
$NOM = 3$.

Then

$$\Delta D_o = 33.6,$$
$$\Delta D_1 = 32.2,$$
$$\Delta D_s = 10.0,$$
$$\Delta D_2 = 33.4,$$

and

$$\begin{aligned} PAG &= 33.6 + 33.2 - 10 - 33.4 - 10\log 3 - 6\text{ dB} \\ &= 11.7\text{ dB} \end{aligned}$$

$$NAG = \Delta D_o - \Delta EAD \tag{34-24}$$

where,

$$\Delta x = -10\log\left(\frac{Q}{4\pi x^2}+\frac{4}{S\hat{a}}\right)$$

where,
x is any distance,
other terms are as before.

Example:
Let
$D_o = 128$ ft,
$EAD = 4$ ft,
$Q = 5$,
$S = 28{,}000$ ft^2,
$a = 0.35$.

then,

$$\Delta D_o = 33.6$$

$$\Delta EAD = 16$$

and,

$$\begin{aligned} NAG &= 33.6 + 16.0 \\ &= 17.7 \text{ dB} \end{aligned}$$

These two equations are mathematically equivalent to Eqs. 34-21 and 34-22 and will produce the same answers given the same data. Eq. 34-21 and 34-22 may be easier to understand and to insert in a computer program.

In addition, it should be noted that some users prefer to place the NOM (number of open microphones) and 6 dB feedback stability margin (FSM) terms in the NAG equation rather than in the PAG equation. This author believes that they belong in the PAG equation since including them produces a value of PAG more nearly equal to that which will be measured in the installed system. While PAG and NAG values will differ with placement of the two terms, the PAG – NAG value (which is the most important result) will be the same regardless of the placement of the two terms.

Also, as before, if PAG is greater than or equal to NAG, it's reasonable to assume that the system will be stable and not feed back.

In Eq. 34-21, the terms D_s, D_1, D_2, D_o, and NOM are as explained in the simplified (outdoor) system. In Eq. 34-22, use the simplified system estimate for EAD Eq. 34-7, ignoring the effects of the reverberant field. This puts the estimate on the safe side for the NAG calculation. The equations for PAG and NAG are similar to the equations given for the simplified system, Eqs. 34-5 and 34-8, except for the 10log[ ] terms that can be interpreted as modifications caused by the room reverberant field.

34.2.3.4.6 The Effect of Directional Microphones and Loudspeakers

In a reverberant room, the effect of directional microphones is less significant than in the outdoor case. The reason is that the amount of reverberant sound energy picked up by the microphone depends very little on the microphone's pickup pattern (a cardioid microphone picks up more reverberant sound from the front, which compensates for its reduced rear pickup). A directional microphone will, however, exhibit higher gain in the direction of the talker which has the same effect on feedback as a reduction in D_s. This improvement may be as much as 2 or 3 dB and it is not included in any of the PAG or NAG equations.

Directional loudspeakers may reduce the amount of direct sound energy reaching the microphone but do not substantially reduce the amount of reverberant sound reaching the microphone, since this is usually dominated by the nondirectional low-frequency loudspeakers. Note that the indoor PAG equation Eq. 34-21 already includes the effect of directional loudspeakers on the reverberant field. Thus, it is best to assume that no additional gain before feedback will be provided by directional loudspeakers.

34.2.3.5 Validity of the Model in a Geometrically Complex Room

In effect, the equations just presented form a mathematical model of the interactions between a room and a sound system. The question arises, "Just how valid is this model?" The answer is to remember that the model assumes a well-developed, statistically random reverberant field in a room with simple geometry. Thus the model can be very accurate in a room like a high-school gymnasium or a rectangular church. Add balconies, transepts, or other complexities, and the equations, while still useful, cannot adequately describe the entire room.

One way to deal with more complex rooms is to treat them as two or more acoustically separate spaces. A large stage with hardwood floors and reflecting walls and ceiling, for example, may be coupled to an audience seating area with padded seats, carpeting, and draped walls. A reverberant cathedral may have an under-balcony area that is very different acoustically from the main room. System design and the use of the equations will be improved by treating these different spaces as entirely different rooms that just happen to share a common boundary (an imaginary wall).

An equation for the overall combined reverberation time in such a dual-space room is

$$RT_{60T} = \sqrt[3]{RT_{60A}^{\;3} + RT_{60B}^{\;3}}\,. \tag{34-25}$$

Some rooms, of course, do not lend themselves to analysis by the given equations. One example is the Superdome-sized room that has no true reverberant field because the individual reflections are spaced far enough apart in time that they are more accurately called echoes, not reverberation. These rooms may have an apparent dramatic increase in reverberation when excited by a high sound pressure level sound system (similar to those used for concert sound reinforcement). Another example is the acoustically dry conference room that has no significant reverberant field because of an abundance of carpeting, draperies, padded seating, and acoustical ceiling tile. One approach to design in both of these spaces is to use the simplified (outdoor system) equations where needed, since they deal with the direct sound. In the large space, echoes must be considered (they are not included in any of the equations); in the small space, table top and other nearby reflections must be considered.

34.2.3.6 A Modification for Low RT_{60} Rooms

The mathematical model presented in this section assumes a well-developed, statistically random reverberant field. Such a reverberant field is unlikely to exist in a corporate conference room or a home living room because of the abundance of sound absorption materials. In many cases, in these small, acoustically dead rooms, the outdoor sound system equations (Eqs. 34-1 through 34-8) can be applied successfully to describe the behavior of sound in the room, Fig. 34-8.

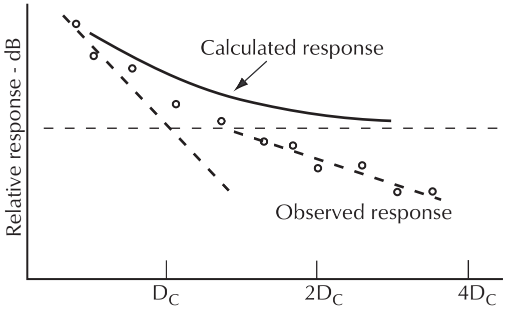


Figure 34-8. Attenuation with distance in a relatively dead room. Courtesy JBL Professional.

Nearby echoes, however, must be considered. These echoes may add to the useful sound at the listeners' ears or they may cause harmful comb filtering depending on the arrival time of these echoes at the listeners' ears compared to the arrival time of the direct sound (see Section 34.2.3.1.2).

When these nearby echoes are useful, they add to the sound level in a way that can be predicted by a modification of Eq. 34-13. Qualitatively, what happens in such a room is that the attenuation of sound with increasing distance from the source is greater than would be predicted from Eq. 34-13. According to Eq. 34-13, the L_P from a source should not attenuate more than 3 dB at any distance past D_c. In these rooms, however, the actual attenuation of sound for distances past D_c is somewhere between the value predicted by Eq. 34-13 and the value that would be predicted by the inverse-square law (Eq. 34-1) as shown in Fig. 34-8. V. M. A. Peutz, one of the originators of the *Alcons* concept, has investigated this phenomenon, and the following equation for attenuation in an acoustically dry room is derived from his work.

$$L_p{}' = L_P - 0.734\left[\frac{\sqrt{V}}{H(RT_{60})}\right]\left[\log\frac{D'}{D}\right] \tag{34-26}$$

where,
H is the room height,
V is the room volume.

Example:
Let
$V = 4275\ \text{ft}^3$,
$H = 9.5$ ft,
$L_P = 90$ dB,
$RT_{60} = 0.4$ s,
$D = 10$ ft,
$D' = 20$ ft.

Note: D and D' must be greater than calculated D_c for Eq. 34-26 to work. D_c calculated for this room = 7.22 ft, so that Eq. 34-26 may be used with $D = 10$ and $D' = 20$.

Then

$$\begin{aligned} L_p{}' &= 90 - 0.734\left[\frac{\sqrt{4275}}{9.5(0.4)}\right]\left[\log\frac{20}{10}\right] \\ &= 86.2\ \text{dB} \end{aligned}$$

Note the similarity between Eq. 34-26 and the inverse-square law Eq. 34-1. The answer, 86.2 dB, is between the 84 dB predicted by Eq. 34-1 and the 88.7 dB predicted by Eq. 34-13. Dimensions are assumed to be in feet. Also, note that Eq. 34-26 should be used only in rooms with very low calculated reverberation times. Eq. 34-26 will accurately predict attenuation for distances D', which are greater than the calculated critical distance, D_c, whenever an on-site measurement shows the actual L_P to be 1 to 5 dB below the predicted L_P at a distance equal to twice the calculated D_c from a source.

34.3 Loudspeaker Systems for Sound Reinforcement

The answer to Question 2: "Can everybody hear?" comes from evaluating the success of the loudspeaker system, in particular, how well the loudspeakers have been aimed to cover the audience and how well the patterns of individual loudspeakers combine to cover areas with complex shapes. To answer Question 2, then, the following section discusses loudspeaker system components, types of loudspeaker systems, and loudspeaker system design.

34.3.1 Loudspeaker Components

A *transducer* is any device that converts one form of energy to another. Loudspeaker components are transducers because they convert electrical energy into acoustic energy. Packaged loudspeaker systems and line arrays are designed from loudspeaker components including cone-type loudspeakers and their enclosures, compression drivers and their horns, and other components such as ribbon drivers and ring radiators. In the past, sound reinforcement systems often used clusters of individual high-frequency horns and low-frequency woofers (cone loudspeakers in enclosures). Today, most systems use packaged loudspeaker systems or line arrays to cover the audience. See Chapter 19 for more detailed information on loudspeakers and Chapter 20 for additional information on cluster design.

34.3.1.1 Cone Loudspeakers

Large cone loudspeakers (15 and 18 inch diameters) are normally used as the low-frequency components of two-way, three-way, or multiway systems. Also, 12 and 10 inch cone loudspeakers may be used as the low-frequency component in a low-power, two-way system or as the lower midrange component in a three-way or multiway system.

Smaller cone loudspeakers (8 and 4 inch) may be used as low-frequency or midrange components in a packaged loudspeaker system. Other 8 and 4 inch cone loudspeakers are designed for relatively full-range performance and are used in ceiling-type distributed systems and as the components in column loudspeaker systems.

34.3.1.2 Cone Loudspeaker Enclosures

There are three basic types of loudspeaker enclosures in use in professional systems: sealed (often improperly called infinite baffle), vented (also called ported or bass reflex), and horn-loaded. Some manufacturers also offer combination vented and horn-loaded enclosures.

A sealed enclosure is relatively simple to design and construct; it has a smooth frequency response curve, good transient response, and helps protect the loudspeaker from overexcursion at low frequencies. Sealed enclosures are most common in home entertainment systems.

A vented enclosure works as a Helmholtz resonator to boost the low-frequency response of a loudspeaker above the response of a similarly sealed enclosure design. Transient response and frequency response smoothness may suffer somewhat, although these problems are small in a good design. An electrical high-pass filter should be used to help protect the loudspeaker against overexcursion at frequencies below the enclosure resonance frequency f_b. Because of their greater output at low frequencies, vented enclosures are common in professional systems.

Horn-loaded enclosures place a horn in front of the loudspeaker and a sealed compression chamber behind the loudspeaker. The loudspeaker thus becomes a compression driver. Properly designed, a horn-loaded enclosure boosts the overall efficiency of the loudspeaker-enclosure combination above a sealed or vented enclosure and provides some measure of control over the dispersion pattern. In addition, the sealed chamber behind the loudspeaker helps prevent overexcursion at low frequencies. Horn-loaded enclosures are most common for midrange applications, Fig. 34-9.

For low-frequency applications, one type of horn-loaded enclosure, often called a vented horn, adds a vented chamber behind the loudspeaker (instead of the sealed chamber) to boost the low-frequency response below the horn's cutoff frequency.

Another type of low-frequency, horn-loaded enclosure, known as a folded horn, is a relatively long horn that has been folded back on itself to reduce the external package size, Fig. 34-10.

Because of their efficiency, horn-loaded enclosures were popular in the early days of sound when power amplifiers were small and expensive. Unfortunately, a horn designed to work well at low frequencies is quite

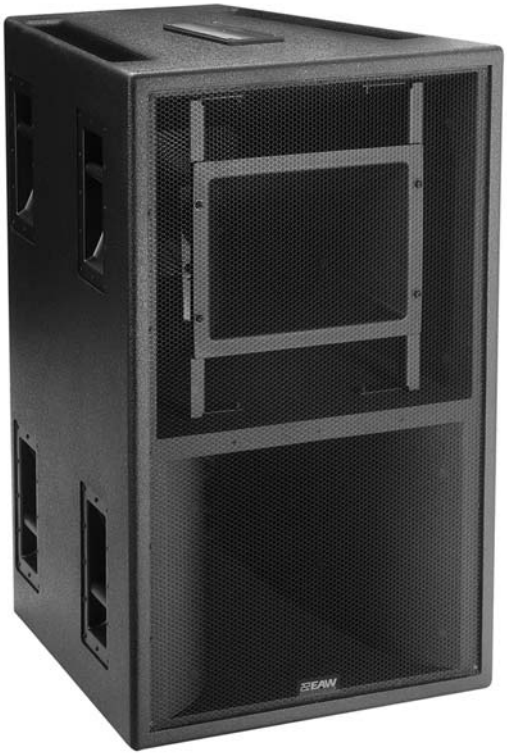


Figure 34-9. A horn-loaded loudspeaker system. Courtesy EAW.

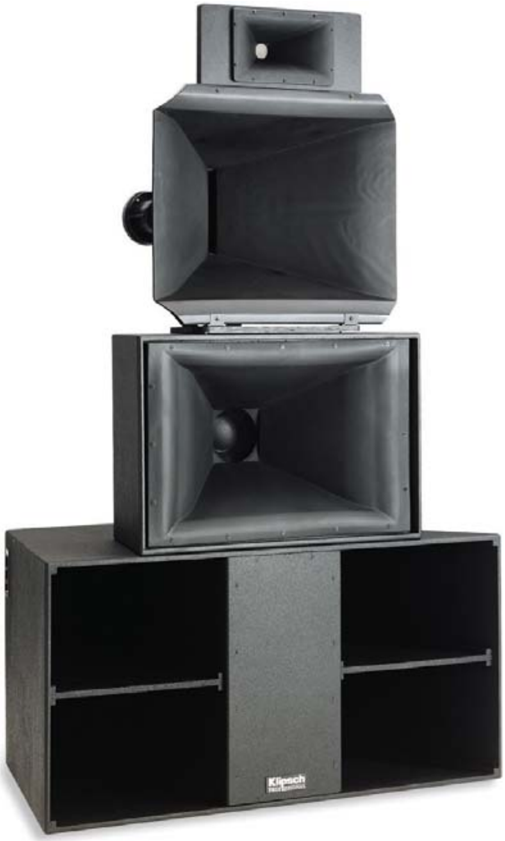


Figure 34-10. A pair of dual 15 inch driver folded-horn enclosures shown with accessory midrange and high-frequency horns. Courtesy Klipsch & Associates, Inc.

large, and with power amplifiers larger and less expensive per watt, many designers now choose vented enclosures because they fit in the smaller spaces provided in modern buildings. It can be shown mathematically that, for a given amount of total enclosure volume, a loudspeaker system using vented enclosures can produce more total acoustic power than a horn-loaded loudspeaker system. The vented system simply uses more loudspeakers and higher-powered amplifiers to achieve this victory.

34.3.1.3 Compression Drivers and Horns

One class of components is designed specifically for use on a horn. These components are called compression drivers, and they are used almost exclusively as the midrange and high-frequency components of two-way, three-way, and multiway systems. At these frequencies, horn sizes are smaller than at the low frequencies used by low-frequency horns. Because of the efficiency and dispersion control of the horn, especially in the critical midrange and high frequencies, mid- and high-frequency horns and compression drivers, Fig. 34-11, are the midrange and high-frequency components most often used in packaged loudspeaker systems.

There are several types of mid- and high-frequency horn designs. These include exponential (radial), multicell, and constant directivity horns. In the past, exponential horns were commonly used in packaged loudspeaker systems and multicell horns were commonly used in component clusters and in cinema loudspeaker systems.

Today, almost all available horns are constant directivity designs. Constant directivity horns have very good dispersion control (pattern control) over a wide frequency band. Although well-designed constant directivity horns are somewhat larger than exponential or multicell horns, the dispersion control advantage of constant directivity horns is so overwhelming that they

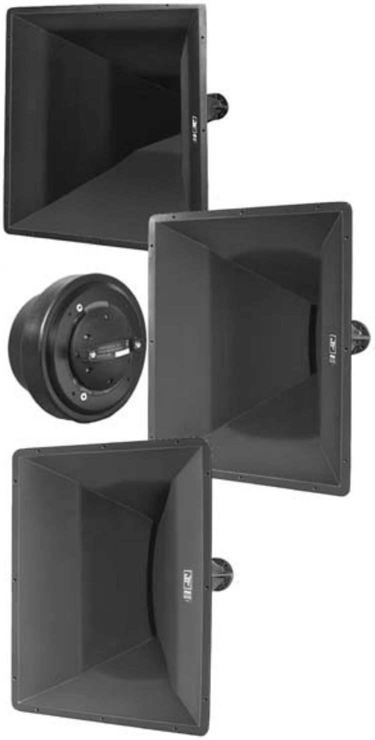


Figure 34-11. A family of constant directivity horns and compression drivers. Courtesy Bosch/Electro-Voice.

have become the most popular choice for packaged loudspeaker systems.

34.3.1.4 Packaged Loudspeaker Systems

Manufacturers now offer a wide variety of packaged loudspeaker systems, designed for many different applications. The overwhelming popularity of packaged loudspeaker systems is due to several factors. First, a packaged loudspeaker combines two or more loudspeaker components in a single enclosure to cover a wider frequency range. As such, a packaged loudspeaker becomes a wide-range component for cluster design. Second, packaged loudspeaker systems commonly include suspension hardware, which makes them easy to install a cluster in the field. Third, packaged loudspeaker systems are usually trapazoidal in shape which makes it easy to arrange them in a tight, efficient cluster. Fourth, packaged loudspeaker systems are generally more attractive than raw components, which is important in appearance-sensitive installations. Finally, some packaged loudspeakers are self-powered and include sophisticated DSP processors that optimize their performance and help protect the loudspeaker. These electronics simplify system design. For all of these reasons, many designers choose packaged loudspeaker systems for their cluster designs, Fig. 34-12.

There are at least two disadvantages to packaged loudspeaker systems. First, in comparison to component horns, packaged loudspeaker systems offer a limited choice of coverage patterns. For example, it's unusual to find a true long-throw horn in a packaged loudspeaker system. Second, designers minimize the size of packaged loudspeaker system for efficient cluster packing and for appearance reasons. However, the smaller horns in these packaged loudspeaker systems have reduced pattern control in comparison to the larger horns available as separate components. In some packaged loudspeaker systems, a three-way design with a midrange horn helps to offset this disadvantage.

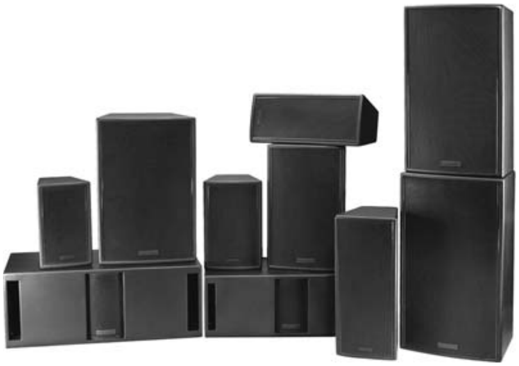


Figure 34-12. A family of packaged loudspeaker systems. Courtesy Community Professional.

At least one manufacturer now offers a packaged loudspeaker system where each model is designed to cover an entire room from a single cluster location. The design achieves this goal by using a specially designed mid/high-frequency horn that has a wide horizontal coverage angle aimed at the front of the room tapering gradually to a narrow coverage angle for the rear of the room. These special-purpose loudspeaker systems work best in rectangular rooms with a specific ratio of length to width. Consult with the manufacturer for additional application advice, Fig. 34-13.

34.3.1.5 Line Arrays

There are two primary styles of line array loudspeaker systems. One is the *modular line array.* also known as a *concert line array.* This type of line array uses multiple

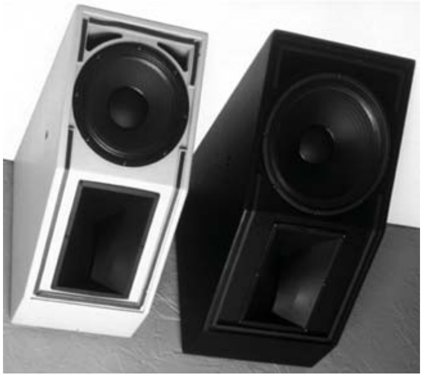


Figure 34-13. EV Vari-Intense. Courtesy Bosch/Electro-Voice.

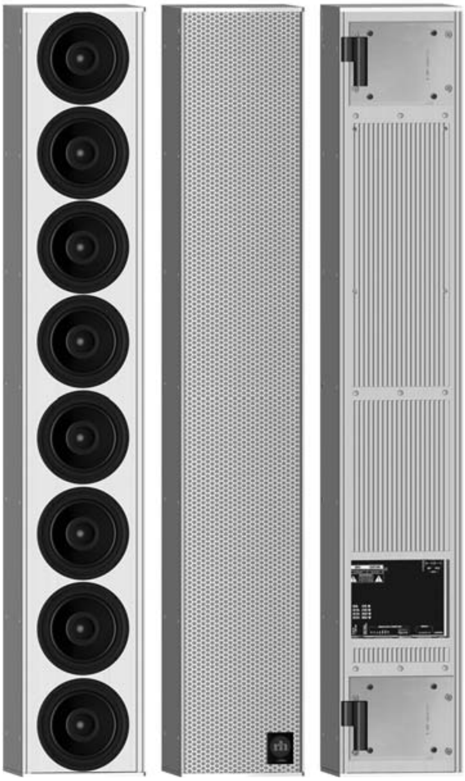


Figure 34-14. Column-style line array. Courtesy Renkus Heinz.

two-way or three-way enclosures suspended in a vertical line. The other type of line array is a *column line array*. This type of line array uses multiple cone-type loudspeakers in a single column-style enclosure, Figs. 34-14 and 34-15.

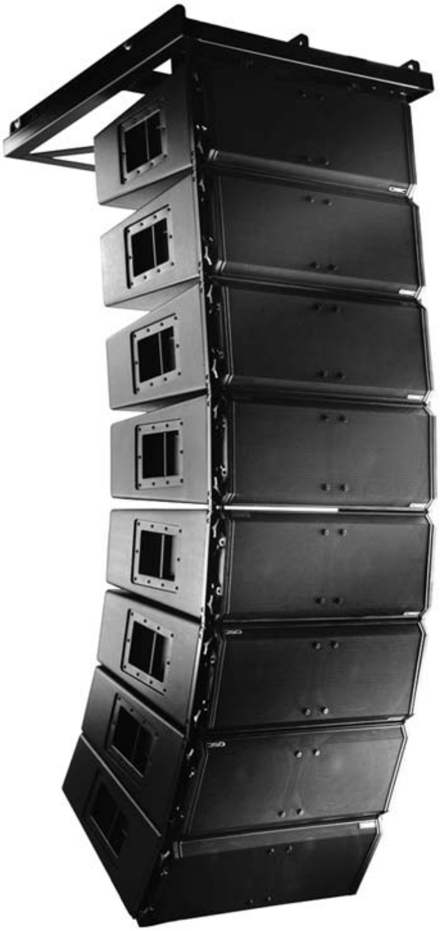


Figure 34-15. Modular line array. Courtesy QSC.

By virtue of their height, both types of line arrays are capable of providing narrow vertical dispersion patterns. This can help the designer keep the sound aimed at the audience and away from a hard reflecting ceiling. Depending on placement it may also help reduce feedback problems by keeping the sound away from system microphones.

Concert line arrays are usually designed with a specific horizontal dispersion pattern provided by horn-type mid- and high-frequency components. Commonly this is a wide angle to cover listeners near the front of the room. In a rectangular room, this wide horizontal dispersion means there will be significant reflections from the walls along the sides of the room. In many rooms, these are beneficial reflections that increase L_p and intelligibility. However, the system designer must be aware of these reflections and confirm that they are useful.

Most column line arrays are designed from cone-type loudspeaker components or ribbon drivers that have very wide horizontal dispersion. As in the concert line-array case, the system designer must confirm that the resulting side-wall reflections will be beneficial for the listeners.

Line arrays offer another benefit to the system designer. For some distance from the line array, the sound level decreases only –3 dB each time the distance from the line array is doubled. Contrast this to normal inverse-square-law loss of –6 dB per doubling of distance. This makes it possible to keep the L_p more constant in an audience area and may also help reduce feedback problems. This effect is frequency dependent and is limited to a distance of about two to two and one half times the height of the array.

Some line arrays include sophisticated electronics and computer software and are steerable. This allows the system designer to aim the vertical dispersion precisely at the audience. Some line arrays even allow two or more lobes that can be aimed at different sections of the audience.

See Chapters 17 and 18 for more information on packaged loudspeaker systems and line arrays.

34.3.1.6 Choosing Loudspeakers

Besides the obvious question of budget, there are several other considerations in choosing loudspeakers that apply to both packaged loudspeaker systems and line arrays.

34.3.1.6.1 Power Handling

The loudspeaker system should be able to handle the expected power output of the chosen power amplifier for an extended period of time over the full-rated frequency range of the loudspeaker.

34.3.1.6.2 Frequency Range and Response

The loudspeaker's response should be smooth over its intended operating range. If the system will be used primarily for voice, a loudspeaker system whose low-frequency response is limited to 70 or 80 Hz should suffice. For music, the system's low-frequency response should extend down to 40 Hz or below. Frequencies below 40 Hz are limited to a few instruments, such as pipe organ, keyboards and synthesizers, and bass drums (kick drum). When it is necessary to reinforce these very low frequencies, use a separate subwoofer system to avoid the added stress these low frequencies would place on the normal system woofers.

34.3.1.6.3 Sensitivity

Sensitivity is an indication of the loudspeaker's efficiency. A loudspeaker's sensitivity is the L_P (sound pressure level) in dB the loudspeaker will produce at one meter, on-axis, when the input power is one watt. High sensitivity is an advantage because it increases maximum L_P. Remember that a decrease of only 3 dB in sensitivity means double the amplifier power is needed to maintain the same L_P.

34.3.1.6.4 Coverage Pattern

Choose the coverage pattern according to the system needs. Packaged loudspeaker systems commonly offer a short-throw (90° × 40°) or medium-throw (60° × 40°) coverage pattern. Some packaged loudspeaker systems may offer wider or narrower coverage patterns. As mentioned, most line arrays have narrow vertical dispersion and wide horizontal dispersion.

Separate horns and woofers are seldom used as the main components in loudspeaker clusters but may be useful for special purposes such as long-throw or balcony coverage or very wide-angle near-throw. Component horns are available in short-throw, medium-throw and long-throw coverage angles. Long-throw component horns are usually 40° horizontal by 20° vertical and are commonly needed only in large concert systems and permanently installed systems. Medium-throw component horns are usually 60° horizontal by 40° vertical and are valuable in many portable as well as permanent systems to reach farther back in an audience. Short-throw component horns are usually 90° or 120° horizontal by 40° vertical and are used to reach the front of an audience or may be used to cover an entire audience in a small portable system.

34.3.1.6.5 Evaluating Loudspeaker Sound Quality

Sound quality is primarily a subjective evaluation, which means that personal tastes play an important part. However, the goal of a sound reinforcement system is not to alter but to reinforce and, to some extent, to enhance the sound of a performance. Thus, the subjective evaluation of the sound quality of a loudspeaker system should be based on how well that loudspeaker system will accurately reinforce a live performance.

For this reason, listening tests done with live sources in acoustically well-designed rooms are ideal evaluations. When it's possible to do live evaluations, use a strong-voice talker with a well-chosen microphone for speech. Use a single, well-known instrument, such as an acoustic guitar or acoustic piano for musical evaluations. A singing voice, accompanied by a guitar or piano is also a good choice. See Chapter 16 for a discussion of microphones.

When a live test is not possible, use a CD or other high-quality digital recording as a source. Certain well-recorded vinyl LPs may also be suitable. Choose a recording of a solo acoustic guitar, acoustic piano, or a voice accompanied by a guitar or piano.

Choosing these simple musical sources makes it easier to evaluate the fidelity of the loudspeaker system because most people are familiar with the way they ought to sound. If the loudspeaker system colors this in any way, most people will recognize the coloration easily.

Using recordings of loud and dynamic rock music or highly synthesized music of any kind may be a good way to evaluate the ability of the loudspeaker system to handle high-power live sources, but it is not a good way to evaluate the fidelity of a system since distortions or frequency response aberrations in the loudspeaker system may be interpreted as intentional parts of the original performance!

34.3.2 Loudspeaker Systems

There are several styles of loudspeaker systems used in sound reinforcement systems. The most common are the central cluster, split cluster, exploded cluster, and the distributed system. Any of these loudspeaker systems may be designed from component loudspeakers, packaged loudspeakers, or line arrays. In addition, there are variations and combinations of these types. This section discusses some basic design criteria for loudspeaker systems. Chapters 17 and 18 discuss additional details of loudspeaker system design.

34.3.2.1 The Central Cluster

A central cluster is a group of packaged loudspeaker systems or component horns and woofers placed in a central location and aimed at a listening area. The traditional central cluster is placed above a stage (on the proscenium) or above the primary microphone location, Fig. 34-16. A modular line array, suspended in this location, may be considered a central cluster.

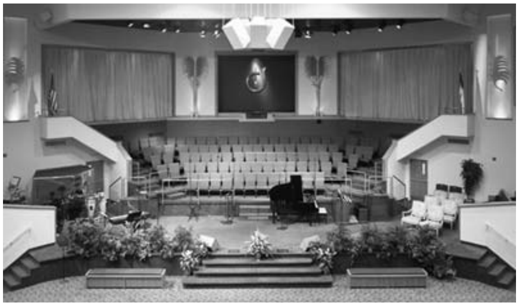


Figure 34-16. Typical central cluster. Courtesy Community Professional.

A location above the audience makes the difference between the distance from the cluster to the nearest listener and from the cluster to the farthest listener more nearly equal. This, in turn, makes the job of designing the cluster for even audience coverage easier. In most cases, however, the cluster should not be more than about 30–45 ft above the heads of the listeners. This is because listeners seated near the talker can often hear both the unaided talker and the cluster. If the cluster is more than about 30–45 ft above the heads of the listeners, they will notice a hollow sound or even a distinct echo due to the natural delay between the sound from the talker and the sound from the cluster.

The human ear can accurately discriminate the location of sounds from a left-right perspective, but not as well from an up-down perspective. Thus, another advantage of a central cluster, if it is placed near the center of the room or approximately above the primary microphone location (and assuming other factors are favorable), is that the sound will appear to emanate from the talker, and not from the cluster.

A final, significant advantage of a central cluster is that, compared to an equally well-designed distributed loudspeaker system, the central cluster is almost always less costly.

Sometimes, aesthetic considerations prevent the installation of a central cluster. For example, the loudspeakers may block important architectural elements in a religious facility or historical building.

In some rooms, the ceiling is too low compared to the length of the room to allow a central cluster to work well. This problem prevents adequate gain before feedback in the back of the room (PAG is too low). A good rule of thumb is that D_2 should be no more than about four times D_1 for a single central cluster to work properly. If the 45 ft height rule is followed, this seems to limit central clusters to rooms with dimensions of 180 ft or less. However, the 45 ft rule can be ignored if the listeners cannot hear the talker without the aid of the sound system. This is often the case in large indoor sports arenas.

34.3.2.2 Variations on the Central Cluster

Sometimes, in a long room with a relatively low ceiling, a second cluster is installed. The second cluster (and the third if more than two clusters are used) is installed some distance out in the room and the sound emanating from these clusters is electronically delayed so that a listener able to hear both clusters will seem to hear only one source, Fig. 34-17. The second cluster effectively divides the room into two rooms, and the first cluster now only has to cover a room that is half as long.

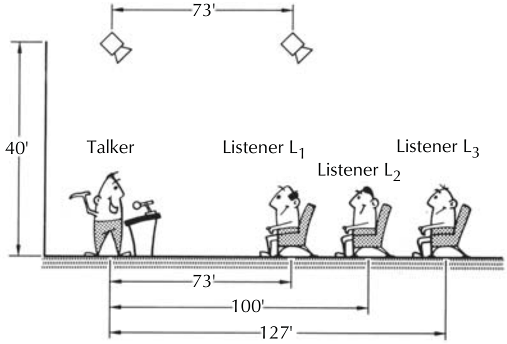


Figure 34-17. Two-cluster (front, back) system. Courtesy Bosch/Electro-Voice.

34.3.2.3 The Split Cluster

One way to install a cluster-type system and preserve central sight lines is to split the cluster with part on the left and part on the right, Fig. 34-18. Unfortunately, this design causes comb filtering in the audience area where the two clusters overlap.

To minimize comb filtering, minimize the amount of overlap between the two clusters. This is easier to accomplish in a room with a central aisle. Another way to minimize this problem, at least partially, is to lower the sound pressure level of one of the clusters about 3 dB and use that lower-level cluster to cover only those listeners who cannot adequately be covered by the other cluster. Design the louder cluster to cover as much of the listening audience as possible.

Line arrays are commonly used in a split-cluster configuration. In particular, column-style line arrays are popular for this application. Because of their narrow vertical dispersion and –3 dB attenuation per doubling of distance, column line arrays can be placed lower than a typical cluster—often at or just above the audience head height. Also, they are perceived as more attractive than other types of loudspeaker systems.

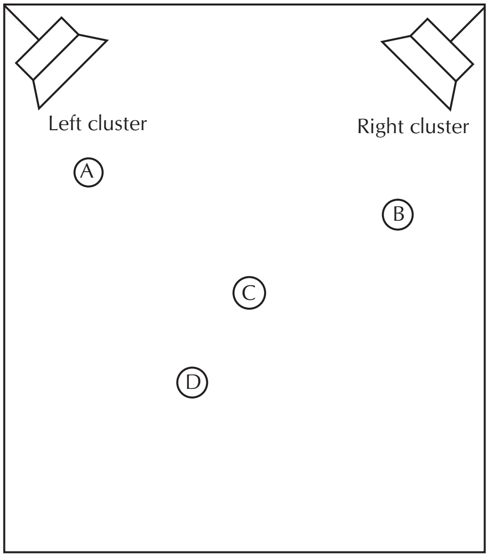


Figure 34-18. The split cluster.

34.3.2.4 Stereo Clusters

For stereo music reinforcement, left and right clusters are required. In order for every listener to hear the stereo effect, both clusters must cover the entire audience. Because this conflicts with the guidelines developed in the previous section, many designers prefer the Left-Center-Right approach discussed next.

34.3.2.5 Left-Center-Right Clusters

For a system that has both voice and stereo or multichannel music, consider a left-center-right cluster design. In this arrangement, the center cluster primarily reinforces the spoken voice and the left and right clusters reinforce the music. It may be difficult for listeners at the left and right edges of the audience to hear the central cluster clearly. In this case, mix a little of the spoken word into the left and right clusters but delay it slightly so the voice appears to come from the central cluster.

It may also be difficult for listeners at the left and right edges of the audience to hear both channels of the music. One way to reduce this problem is to mix some of both the left and right channels of music into the central cluster. Of course, this will reduce the stereo effect for listeners seated in the center of the audience.

Left-center-right clusters are commonly used for dramatic performances where actors' voices are panned from one side of the stage to the other. In this type of system it is imperative that every listener be able to hear all three clusters.

34.3.2.6 Exploded Clusters

Technically, the term *exploded cluster* refers to a conventional central cluster where individual packaged loudspeakers have been moved outwards (exploded) along radii from the original position. The term is often used, however, to describe any system where several, smaller clusters are used in place of a single, central cluster. In this sense, a left-center-right cluster system could be considered an exploded cluster.

Exploded clusters are often used in auditoriums with very wide stages and relatively shallow audience areas, Fig. 34-19. Many so-called mega churches in the United States are of this design. A central cluster cannot cover the sides of the room toward the stage effectively. A split cluster (including line arrays) cannot cover the center of the room toward the stage effectively. An exploded cluster is often a good compromise for this type of facility. Typically, left, center, and right clusters will be supplemented by an additional cluster between the left and center clusters and another between the right and center clusters.

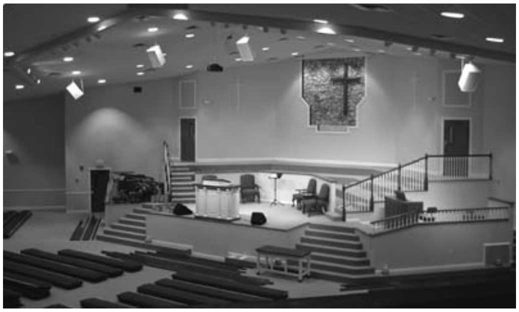


Figure 34-19. An exploded cluster loudspeaker system. Courtesy Community Professional.

To minimize comb filtering, try to design each cluster to cover a specific audience area and minimize overlap. For example, if there are four audience areas, separated by aisles, try to design the system with four clusters. If the room is fairly deep, consider a second ring of delayed clusters, located outward on imaginary radii from the near-stage clusters. An exploded cluster system can also be combined with delayed under-balcony speakers.

34.3.2.7 Rear and Surround Clusters

Rear and surround clusters are primarily used for special effects in live dramatic presentations. Design these clusters for the facility and presentation requirements. For example, if a rear cluster is to reinforce the spoken voice of an off-stage actor, it must cover the entire audience clearly. However, left and right clusters used for ambience effects may not need to cover the entire audience evenly.

34.3.2.8 Designing a Central Cluster in a Simple Rectangular Room

Most sound reinforcement system designers now use EASE or some other loudspeaker system design software to help them design their systems. The following method, used before this type of software was available, is presented for educational purposes and for historical completeness. In addition, some of the concepts presented here, such as methods of choosing loudspeakers, will be of value to a designer using EASE.

34.3.2.8.1 Evaluate the Room

If the room exists, measure its reverberation time and physical dimensions (drawings will help in physical measurements). Calculate the room volume and total surface area. Using the Sabine reverberation time equation, derive the average absorption coefficient, $\bar{a}$, as follows:

$$RT_{60} = \frac{0.049V}{S\bar{a}} \tag{34-27}$$

where,

V is the room volume,

S is the room surface,

$\bar{a}$ is the average absorption coefficient,

For metric distances, replace the constant 0.049 with the constant 0.161.

From this equation, the average absorption coefficient, $\bar{a}$, can be found with

$$\bar{a} = \frac{0.049V}{S(RT_{60})} \tag{34-28}$$

For metric distances, replace the constant 0.049 with the constant 0.161.

If the room is in the planning stages only, estimate its average absorption coefficient, total surface area, and volume from the architectural data.

34.3.2.8.2 Choosing the Cluster Location

The ideal cluster location will probably be approximately above the primary microphone location; that is, in a rectangular room, the cluster should be near the top and at the center of an end wall. Compromise locations are discussed in the following sections.

34.3.2.8.3 Evaluating the Cluster Location

The potential success of a proposed cluster location and cluster design can be evaluated by answering "The Four Questions." starting with Question 4: "Will it feed back?" Using available data and Eq. 34-21, answer this question before moving on to choose loudspeaker types.

34.3.2.8.4 Choosing the Loudspeakers

The following discussions assume the design will center around constant-directivity packaged loudspeaker systems for a speech or speech and music reinforcement system.

Begin the design by choosing a loudspeaker for coverage in the rear of the room. Use trigonometry or one of the cluster layout methods discussed in Section 34.3.2.9, and determine the horizontal (side-to-side) coverage angle required at the rear of the room. Using more than one loudspeaker to cover an area is discussed in the next section. Remember that the listener's ears are approximately 4 ft (1.2 m) above the floor when determining required coverage angles. (This can be simulated in most design methods by placing an imaginary floor at 4 ft [1.2 m] above the actual floor.)

Once a rear coverage loudspeaker has been chosen, use its Q value (from the manufacturer's specification sheet) and the known parameters of the room to answer Question 3: "Can everyone understand?" by calculating the articulation loss of consonants (*Alcons*) from Eq. 34-18. Assume $N = 1$, then do the calculation over again using $N = 2$ and $N = 3$ to simulate the effects of adding loudspeakers to the system. If *Alcons* is less than or equal to 10% for each of these calculations, the system design will work from the criteria of Question 3. If *Alcons* exceeds 10% for any of these calculations, a second cluster or a distributed system may be required. (Note that some designers believe an *Alcons* of 15% or less is acceptable.) It might seem logical, in this case, to simply choose a loudspeaker with a higher Q value, since this would reduce the *Alcons*. That loudspeaker will also have a narrower coverage pattern, however, and might not adequately cover the entire audience. Using additional high Q loudspeaker will not solve the problem either since they increase the value of N. Thus, if *Alcons* is too high, about the only alternative to a second cluster or distributed system is to reduce the room reverberation time with acoustic treatment.

34.3.2.8.5 Aiming the Loudspeakers

If the *Alcons* is acceptable, choose additional loudspeaker to cover the rest of the room, providing a yes answer to Question 2: "Can everybody hear?" In many rectangular rooms, only one or two additional (wider angle) loudspeakers will be required.

The edge of the defined coverage pattern of a loudspeaker is its –6 dB point. For this reason, it is common practice to overlap the coverage patterns of the loudspeakers in a cluster to compensate for the drop-off in coverage of each individual loudspeaker near the edges of its coverage pattern. A good rule for this practice is to overlap two loudspeakers as little as possible to maintain consistent coverage throughout an audience area. Another good rule is to aim the overlap areas at an unimportant part of the audience area such as an aisle.

Avoid aiming loudspeakers at hard rear or side walls (which could result in echoes) and avoid aiming them directly down at the microphone (which could increase the possibility of feedback). Again, remember to aim the loudspeakers at the listener's ears (about 4 ft [1.2 m]above the floor for seated listeners or 5 ft [1.5 m] for standing listeners).

Now, re-evaluate Question 4: "Will it feed back?" by performing the calculations for PAG and NAG discussed in Section 34.2.3.4.5. The answers to the PAG and NAG equations may be tempered by recalculating D_c for an increased value of N. If feedback seems possible, consider moving the cluster, or consider using an automatic microphone mixer to keep NOM = 1 or, when possible, teach the users to talk closer to the microphone (which reduces D_s), see Section 34.4.3. Using directional loudspeakers and microphones may provide some additional gain before feedback, but it's best not to plan for this additional gain.

34.3.2.8.6 Powering the Cluster

The remaining question, Question 1: "Is the system loud enough?" can be answered by choosing loudspeakers and power amplifiers to satisfy Eq. 34-16 remembering the criteria for head room and SNR.

34.3.2.8.7 Low-Frequency Loudspeakers and Feedback

The published Q for a two-way packaged loudspeaker system is a compromise between the relatively high Q of the high-frequency horn and the lower Q of the woofer. The same is true of three-way and four-way packaged loudspeaker systems. Using this published Q in the PAG and NAG equations may provide an acceptable result. However, the lower Q value of the woofer may, in some systems, result in feedback in the woofer's frequency range.

34.3.2.8.8 Direct Sound and Feedback

The PAG and NAG equations assume the microphone is entirely in the reverberant field of the loudspeaker cluster. If the microphone is receiving any significant amount of direct sound from either low- or high-frequency loudspeakers, the PAG and NAG equations will not accurately predict feedback problems. This potential direct sound feedback problem must be considered qualitatively in the design of the cluster by aiming the loudspeakers away from the microphones. Since the low-frequency loudspeakers are normally low Q and cannot be effectively aimed away from the microphone, they are often placed nearest the ceiling to be as far from the microphone as physically possible.

34.3.2.8.9 Other Considerations

At this point, the basic cluster design is finished. Consideration should be given, of course, to overall system head room, frequency response, distortion, and other sound-quality factors that help to answer Question 5, "Does it sound good?"

34.3.2.9 Answering Question 2: "Can Everybody Hear"

The cluster design process outlined in the last section includes little help in actually choosing and aiming the horns. Trigonometric methods may suffice for a simple room. For more complex rooms, most designers use EASE or Modeler or another computer software design tool.

34.3.2.9.1 History of Manual/Graphical Cluster Design Tools

Before the advent of EASE, Modeler, or other software design tools, system designers used one of several manual/graphical tools that were extensions of architectural mapping techniques. A user of one of these tools would overlay a graphical representation of the loudspeaker coverage pattern, known as an isobar. onto a specially prepared angular map of the room.

Because both the room map and loudspeaker isobar were angular, moving the isobar around the room map, was equivalent to re-aiming the horn. Thus, the user could aim the loudspeaker to estimate optimum coverage patterns. More than one loudspeaker isobar overlay could be used to estimate the coverage patterns of multiple, overlapping horns.

The concepts behind these methods were developed by several engineers including Thomas McCarthy of North Star Sound, and first commercialized as Altec's Array Perspective. Fig. 34-20. The disadvantage of the Altec method was its lack of accuracy, especially when horn aiming angles were far from the zero line defined for the room map.

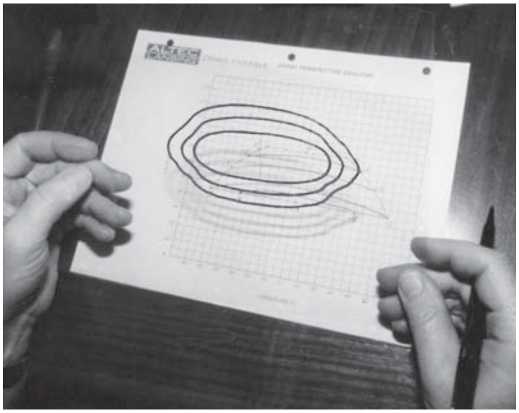


Figure 34-20. Array Perspective, a manual/graphical design method. Courtesy Bosch/Electro-Voice.

Another manual/graphical method, developed by John Prohs and David Harris at Ambassador College, used a transparent plastic sphere onto which the room was mapped. Semispherical plastic overlays represented horn patterns and could be moved around on the room map to simulate aiming the horns. This method resolved most of the accuracy problems of the flat transparency method and was commercialized by Community Light and Sound as the "Cluster Computer."

A number of noted engineers and consultants contributed to the development of these manual/graphical design tools. Among them were Ed Seeley, Thomas McCarthy, Ted Uzzle, John Prohs and David Harris, Farrel Becker, Peter Tappan, Gene Patronis, and Bob Thurmond.

34.3.2.9.2 Commercial Computer Software Design Tools

As personal computers grew in power, the manual/graphical design tools were ported to the PC. Among the early computer programs for cluster design were JBL's CADP, Bose Modeler, Electro-Voice/Altec AcoustaCADD, Thomas McCarthy's Umbulus, and John Prohs PHD Program.

The cost of maintaining these programs and compiling the necessary database of loudspeakers is high and two commercial software programs have come to dominate the field. The first, EASE, was developed by consultant Dr. Wolfgang Anhert, and is distributed by Renkus-Heinz, Fig. 34-21. The other, Modeler, is a proprietary program developed and distributed by Bose. EASE runs on a Windows-based PC. Modeler runs on an Apple Macintosh or a Windows-based PC.

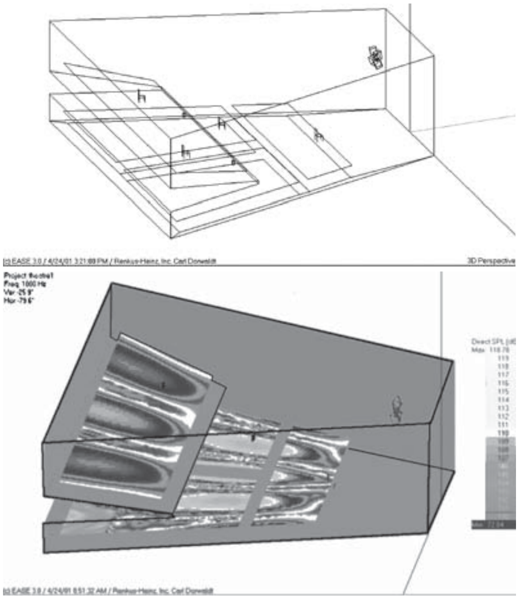


Figure 34-21. EASE. Courtesy Renkus-Heinz.

Modeler, while a very capable program, is only usable with Bose loudspeakers. Check with the Bose Corporation for updates to this policy. As a result, EASE, which can be used with any manufacturer's loudspeakers, is preferred by most designers. Either program has a considerable learning curve and designers must attend one of the available training classes.

Before designing a cluster with either program, the user must enter a detailed room description in three dimensions. This means the user must have accurate and detailed room dimensions and should be familiar with computer aided drafting methods. The programs will import properly prepared computer aided drafting files, such as those produced with AutoCAD.

To use a chosen loudspeaker in either program, the loudspeaker manufacturer must provide compatible data files. Most major loudspeaker manufacturers provide these files but data files for certain loudspeaker models may not be available.

The requirement to enter detailed room data and the lack of availability of some loudspeaker data are limitations that may inhibit a designer from using EASE or Modeler on certain projects. However, the power and versatility of these programs are very high and most designers will want to utilize one of these programs for most projects. Chapters 9 and 35 provide additional details about computer room modeling and auralization.

34.3.2.9.3 Other Software Tools

Additional software tools are available for specialized applications. As an example, Syn-Aud-Con offers a multifunction spreadsheet that calculates many of the equations in this chapter and provides other useful functions.

34.3.2.10 Designing the Complex Cluster

Although programs like EASE and Modeler handle complex clusters well, it's important for the designer to understand the difficulties behind such a design. In concept, the design of a multiloudspeaker central cluster in a room with complex geometry is no different from the simple rectangular room design discussed in Section 34.3.2. In practice, the complex cluster shown in Fig. 34-22 presents a set of new difficulties.

First, the complex cluster is most often designed for a large public facility. In practically every such situation, the budget will be tight, leaving no room for errors in design. It is simply not economically possible to redesign a large cluster after it has been installed.

Second, many of the approximations used in the simple cluster design method are too gross to be used in the design of a complex cluster. For example, the approximations used for the value of N in the *Alcons* equation need to be refined for a complex cluster design.

Here, based on the four questions, is a primarily qualitative explanation of some of the refinements needed for the design of a complex cluster. Some quantitative explanations are given, but a full, quantitative analysis of the complex cluster is beyond the scope of this discussion. EASE or Modeler will greatly help in

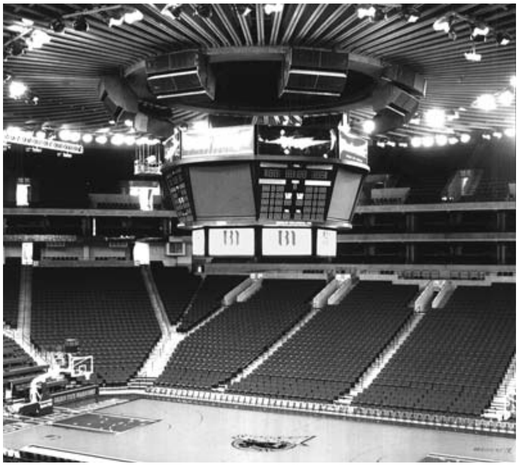


Figure 34-22. A large, complex cluster. Courtesy JBL Professional.

the design of a complex cluster, but *only when all room and loudspeaker data is accurate*. In any case, the complexity of the design process and the economic consequences of errors are significant enough that the services of an experienced, qualified acoustical consultant are highly recommended.

34.3.2.10.1 The Fundamental Complexity

When a cluster involves more than perhaps two or three loudspeakers, its operation becomes complex. The calculation of *Alcons* for a far-throw loudspeaker, for example, cannot ignore the reverberant field contributions from the other loudspeakers in the cluster. A qualitative method of dealing with the problem is straightforward. The direct sound level at each listener can be calculated via the inverse-square law, Eq. 34-1, from a knowledge of which loudspeaker or loudspeakers are aimed at the listener. The total reverberant field sound pressure level can be calculated by adding the total acoustic power output of all the loudspeakers and placing this value into a modified form of the room reverberation equation. Answering the four questions becomes a matter of using either the direct sound, the direct plus reverberant sound, or a direct/reverberant ratio.[5,6]

Calculating the reverberant field level in the room requires a detailed knowledge of the room's acoustic parameters, however, and in many rooms, the acoustic parameters change from position to position in the room. A religious facility in the shape of a cross (cruciform church), with balconies in each wing, for example, may have several totally acoustically different spaces that can be covered from the same cluster location. While each space has different acoustics, they interact with each other in a complex way further complicating the process of calculating the reverberant field at any listener's location, Fig. 34-23.

Even in rooms that are fairly well behaved acoustically and have a statistically random reverberant field, the calculation of reverberant field level is not simple. This is because this calculation requires a thorough knowledge of the characteristics of the cluster. Each loudspeaker adds an amount to the reverberant field that depends on the electrical power applied to the loudspeaker and its efficiency. While few manufacturers provide direct data on the efficiency of their loudspeakers, this can be calculated from a knowledge of the on-axis sensitivity and the Q.

Finally, all the factors involved in the reverberant field calculation vary with frequency so that, for full precision, a new set of calculations must be performed for each frequency of interest.

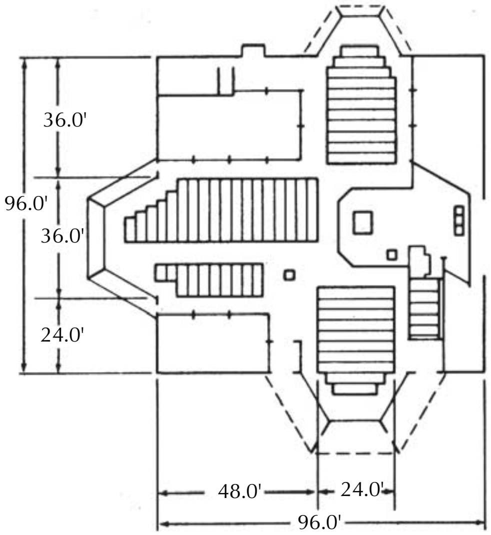
34.3.2.10.2 Simplifying the Complex Cluster Design Process

Before the advent of EASE, Modeler and the earlier software design tools, these complexities were so overwhelming that few designers made any attempt to deal with them directly. Instead, they used simplifications such as the D_c Modifier approach. This approach involves calculating a value of N for the cluster.

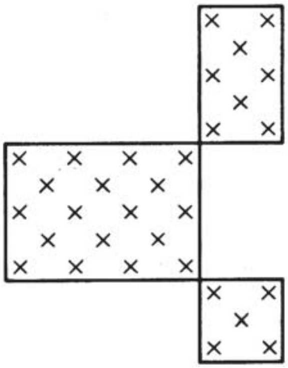
A simple estimate of N is the total number of loudspeakers producing the same acoustic power as the loudspeaker pointed at any given listener. For example, if there are three loudspeakers in the system, all producing the same acoustic power, and only one loudspeaker is pointed at the listener, then N equals 3. If there are four loudspeakers, two of which are pointed at the listener, the situation is the same as having two loudspeakers with one pointed at the listener. Therefore, N would equal 2.

It's possible to extend this simple estimate in a logical manner resulting in a value of N for almost any cluster. However, the process itself is complex and it depends on estimates for the acoustic power output of each loudspeaker calculated from the loudspeaker's published Q and L_s, both of which are compromise values that vary with frequency.

A better way is to rely on a software evaluation of the complex cluster performed by EASE or Modeler or one of the other software design tools. The detailed loudspeaker data required by these software tools enables them to calculate the acoustic power output of each loudspeaker and, thus, its contribution to the rever-



A. Cruciform church, loudspeaker coverage required over seating area only.



B. Scaled drawing of the seating area for the cruciform church using 4-inch loudspeakers at 4000 Hz.

Figure 34-23. A cruciform church showing distributed loudspeaker coverage. Courtesy Bosch/Electro-Voice.

berant field. Given the direct and reverberant field level at a listener's position, the answers to Questions 2 and 3 are simple to calculate. Because the software tools do this so well, few designers attempt to design a complex cluster without using them.

34.3.2.10.3 The Four Questions for the Complex Cluster

34.3.2.10.4 Question 1: Is the System Loud Enough?

For the complex cluster, a new approach is indicated. That approach is to find the total direct sound level at the listener's position, then to find, independently, the total reverberant sound level at the listener's position, and to add them to get the overall total sound level at that position. Comparing and adjusting this overall level with the desired L_P (i.e., is it at least 15 dB above ambient noise?) will provide the required EPR. It is also possible to use the original indoor EPR equation (Eq. 34-16) by calculating a value for the D_c modifier N and using this value of N in the critical distance portion of the EPR calculations.

Many designers believe that, although the reverberant energy in a room can aid the perception of useful loudness, it is the direct sound level that is most important. The L_{Pd} can, of course, always be calculated via the inverse-square law, Eq. 34-1. If a listener is in the direct field of more than one loudspeaker, the direct sound from all such loudspeakers can be added to obtain the total L_{Pd} at the listener. If the L_{Pd} level at the listener is high enough, the L_{Pt} will also be high enough, thus, answering Question 1 satisfactorily.

34.3.2.10.5 Question 2: Can Everybody Hear?

For a complex cluster, the answer to Question 2 can be obtained manually. However, given accurate data, software tools like EASE and Modeler provide a detailed and accurate answer that makes these tools the best way to answer Question 2.

34.3.2.10.6 Question 3: Can Everybody Understand?

The *Alcons* equation, like the EPR equation, may be used for a complex cluster by calculating a value of N as discussed in Section 34.3.2.10.

Another method is to calculate the direct field and the reverberant field at each listener's position of interest and perform a direct/reverberant comparison. Since the *Alcons* concept is based on the direct/reverberant ratio, a knowledge of the actual numeric direct/reverberant ratio is equivalent to a knowledge of the numeric value of *Alcons*. However, the *Alcons* equation assumes a 25 dB SNR. In rooms with a higher level of ambient noise, use Eq. 34-20. EASE and Modeler also provide good estimates of intelligibility at each listener's position.

34.3.2.10.7 Question 4: Will It Feed Back?

This question can be answered by calculating the total sound level (direct plus reverberant) reaching the microphone from the cluster. If this level is equal to or greater than the level expected from the talker at the microphone, then the system will almost certainly feed back. If the total direct plus reverberant level at the microphone is at least –6 dB below the expected level from the talker, the system will probably be stable.

Reference 6 provides a method, if complicated, for calculating the total reverberant level at the microphone from a complex cluster. The PAG and NAG equations may be used with a value of *N*, calculated as discussed in Section 34.3.2.10 for the complex cluster, although the direct plus reverberant approach will probably give a more accurate result. EASE and Modeler also provide this capability.

The actual feedback process also involves reflections from near the microphone, since these reflections can increase the sound level at the microphone by as much as 6 dB (for a reflection from a hard surface that is in phase with the source sound and, therefore, adds coherently with the source sound). Thus, such a reflection could cause a system to feed back even when calculations would show a 6 dB feedback stability margin.

Addition of more than one microphone to the complex cluster system will not automatically lower the gain before feedback by 3 dB as indicated in the PAG equation. That is because each microphone may receive a different amount of direct sound from the cluster and may be subjected to different nearby reflections. Assuming similar microphone locations, however, simplifies the calculation, and a 3 dB reduction in gain before feedback is a reasonable assumption in most systems. This means that the total sound level from the cluster reaching either microphone must be about –9 dB below the expected talker level at that microphone to achieve the 6 dB feedback stability margin. This also, of course, assumes that the gain (volume control setting) of each microphone is similar.

34.3.2.11 Signal Alignment and Purposeful Misalignment in Cluster Design

Ideally, the sounds from two loudspeakers that cover the same audience area reach the listeners' ears at precisely the same moment. However, experienced designers know this is almost impossible to achieve in a real system.

One proposed solution to this problem is signal alignment of the cluster. To implement this concept, the system must be designed with a separate amplifier channel, and separate signal delay channel for each loudspeaker. The signal delay must be adjustable in 10 μs (or smaller) increments. Choose a reference loudspeaker, usually a long-throw loudspeaker, and delay all the other loudspeakers so their signals line up in time with the reference loudspeaker.

Signal alignment can work well when the audience is concentrated in a small area. However, it is very difficult to design a cluster to be properly signal aligned for a larger audience area. Aligning the cluster for one listener may make matters worse for another listener. Also, correct signal alignment can vary with frequency. As a result of these and other problems, signal alignment is best for very simple clusters or for clusters that have been specifically designed for signal alignment.

Another solution to the problem of signal alignment is to purposely misalign the loudspeakers in such a way as to simulate beneficial reflections (see Section 34.2.3.1.2). This concept relies on the idea that a delayed signal arriving between approximately 20 ms and 50 ms after the original signal is beneficial to both L_p and intelligibility.

Purposeful misalignment ideas can be used to advantage in designing an exploded cluster. Following this concept, it is acceptable for a listener to hear two clusters at once if the signal from the second cluster arrives between 20 ms and 50 ms after the signal from the first cluster. Some designers use DSP signal delay on alternate clusters in an exploded cluster to achieve this goal.

It is even possible to combine the concepts of signal alignment and purposeful misalignment by carefully aligning a small cluster and purposely misaligning that cluster with another, nearby cluster. However, there is as much art as science in the implementation of either signal alignment or purposeful misalignment. For this reason, the designer is cautioned to rely on traditional cluster design principles and to not depend on signal alignment or purposeful misalignment for the success of a design.

34.3.3 The Distributed Loudspeaker System

There are several types of distributed loudspeaker systems but all share a common theme. In contrast to a central cluster where the loudspeakers are all concentrated in one location, the loudspeakers in a distributed system are distributed throughout the audience area in such a way as to cover the area evenly.

Because every listener is more or less the same distance from a loudspeaker, coverage can be very uniform from a distributed system. In addition, if the system is carefully designed, the potential for feedback should be very low, and because every listener is relatively near a loudspeaker, the direct/reverberant ratio is high and intelligibility is often very good.

Thus, in the ideal case, a well-designed distributed system can work very well. One problem with even the ideal distributed system, however, is that the listeners will hear the sound coming from over their heads. That is, the natural localization provided by a central cluster is not provided by a distributed system. This problem is usually minor and can be minimized by using an electronic signal delay in combination with a localizer loudspeaker, a subject covered in Section 34.3.3.8.1.

Sometimes, a distributed system will be installed in a very large room with a high reverberation time. A convention hall exhibit room or high-ceiling airport terminal is a good example. This type of system will probably be used for paging and perhaps for background music. Thus, the localization of a central cluster is not needed. The high reverberation and high noise present in rooms like this, however, present problems to the system designer. These problems can be at least partially overcome by a dense enough layout and careful equalization. In these areas *N* also plays an important role. The distributed system has an additional advantage over the central cluster in this case because it effectively reduces the value of D_2 (the distance from the loudspeaker to the farthest listener) and, thus, improves the SNR and the direct/reverberant ratio.

Another reason for using a distributed system, even in a room that could utilize a central cluster, is that the distributed system allows a more flexible room layout than the central cluster. In some multipurpose rooms, for example, the stage location may be changed from event to event, and some events may not use a stage at all. The distributed system allows almost any location to be used as the stage or primary microphone location without the distraction that would be caused by having a central cluster behind the heads of the audience. In addition, in large, reverberant spaces, like sports arenas, distributed loudspeakers above unused portions of the room can be turned off. This helps intelligibility because it improves the direct/reverberant ratio by lowering the amount of energy uselessly put into the reverberant field.

The primary disadvantage of any distributed system is its (usual) higher cost compared with a central cluster designed for the same space (assuming that a central cluster could work in the space).

34.3.3.1 Distributed Ceiling Loudspeaker Systems

Distributed ceiling loudspeaker systems are normally installed in rooms with low ceilings (low compared to the length and width dimensions of the room) where a central cluster or distributed cluster system cannot adequately cover the room. In some situations where a central cluster would work from a design point of view, a distributed system is chosen for its versatility. A distributed system (without delay) does not have the psychoacoustic localization of a central cluster. Thus, microphone locations can be varied without worrying about the effects on localization. Loudspeakers can be turned off above microphone locations making multiple (or varying) microphone locations possible while reducing feedback potential. Loudspeakers can also be turned off in areas not in use to avoid exciting the reverberant field in those areas. Adding variable signal delay makes it possible to provide the psychoacoustic localization of a central cluster from any chosen microphone location, Fig. 34-24.

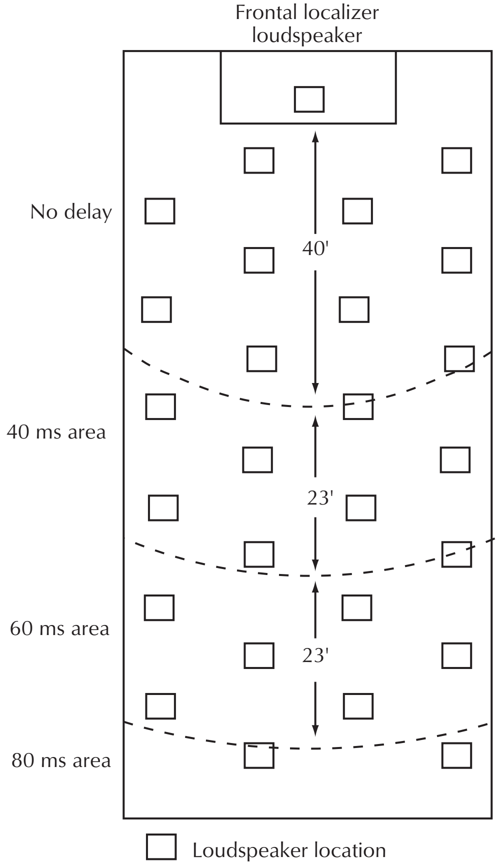


Figure 34-24. Ceiling distributed loudspeaker system with signal delay and localizer loudspeaker. Courtesy Bosch/Electro-Voice.

34.3.3.2 Central Cluster Plus Distributed System

In some rooms with central clusters and rear balconies, listeners seated under the balcony are not adequately covered by the central cluster. In this case, a distributed system, placed under the balcony, with signal delay, can provide coverage for the under-balcony listeners, Fig. 34-25. A good rule of thumb is that every listener under a balcony should have good line of sight to the loudspeaker covering his or her area in order to hear the central cluster well. If any part of the central cluster is shadowed by the balcony, consider a distributed, under-balcony, signal-delayed system.

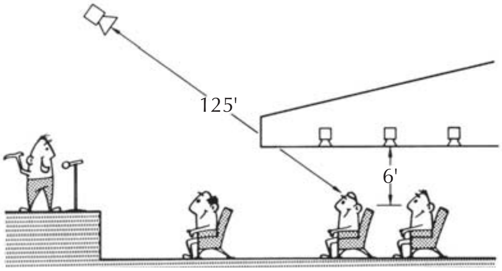


Figure 34-25. Central cluster plus under-balcony distributed loudspeakers on signal delay. Courtesy Bosch/Electro-Voice.

34.3.3.3 Multicluster System

A multicluster system could be considered a distributed system. Also, in the large arenas and exhibit halls mentioned previously, horn/woofer components more typical of a central cluster may be used in the distributed system simply to produce a higher L_P capability.

34.3.3.4 Distributed Column System

One common variation on the distributed system is called a distributed column system, Fig. 34-26. This type of system is normally installed in a long, narrow religious cathedral where a central cluster cannot be installed for aesthetic reasons. A group of column-type loudspeakers (or some other type of packaged loudspeaker systems with appropriate dispersion) are installed. Electronic signal delay and a localizer loudspeaker may be included. This system can have the same problems as the split cluster described earlier, however, and should be used only in narrow rooms to minimize these problems.

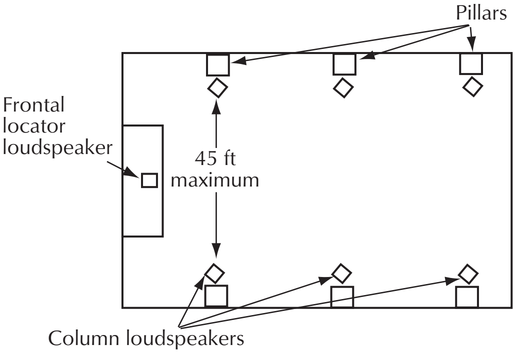


Figure 34-26. A distributed column loudspeaker system with signal delay. Courtesy Bosch/Electro-Voice.

34.3.3.5 Pew-Back Distributed System

Another distributed variation for churches is known as the pew-back system. In this system, pioneered by consultant David Klepper, a large number of small loudspeakers (usually one loudspeaker for every two to three listeners) is placed on the backs of the pews and facing the listeners. Signal delay is required. One problem with a pew-back system is the significant change in D_2 and loudspeaker Q when the listeners stand up (the listeners move both farther away from the loudspeaker and more off axis of the loudspeaker). Another problem is the large value of N caused by the large number of active loudspeakers. To partially overcome the N problem, add a latching relay in each enclosure and a "push-to-listen" button. This way, loudspeakers in pews with no listeners will not be turned on, and those loudspeakers will not uselessly add to the reverberant field intensity.

Both the distributed column system and the pew-back system are difficult to design and install properly. As with any difficult design, an experienced acoustical consultant may be the best answer to getting a costly job done right the first time.

34.3.3.6 Loudspeakers for Distributed Systems

34.3.3.6.1 Ceiling Loudspeaker Systems

Typical ceiling loudspeakers are 4 inch or 8 inch cone-type components that often come in a package that includes a round metal enclosure, grille, and 70 V transformer, Fig. 34-27. Those designed for installation in a dropped ceiling may also include an optional T-bar suspension system. Some may be UL listed for fire signaling or fire resistance. This type of ceiling loudspeaker is intended for sound reinforcement, business music, and paging in applications that do not require high SPL levels.

Larger ceiling loudspeakers and enclosures are designed for use in convention centers, arenas, and other applications that may require mid- to high SPL levels and a wide frequency range. These systems typi-

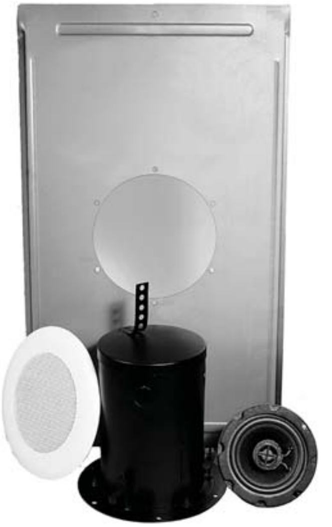


Figure 34-27. Typical ceiling loudspeaker components. Courtesy Atlas Sound.

cally use a 12 or 15 inch coaxial loudspeaker component, Fig. 34-28.

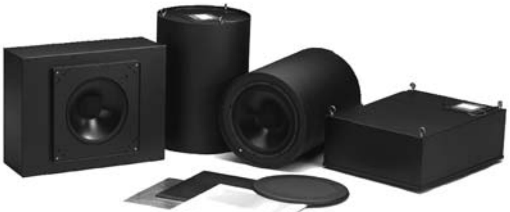


Figure 34-28. Large ceiling-type loudspeaker systems. Courtesy Lowell.

34.3.3.6.2 Packaged Loudspeaker Systems in Distributed Systems

Packaged loudspeaker systems are commonly used in distributed systems in large rooms with relatively low ceilings. Examples include convention centers, exhibit halls, and hotel ballrooms, Fig. 34-29.

In these rooms, small, packaged systems may be hung, face down, in a distributed fashion to cover the room. The advantages of high-quality packaged systems in this application are their high output level and their wider range frequency response in comparison to typical 8 or even 12 inch ceiling-type loudspeakers. Rigging safety and fire safety must be considered for these applications.

Packaged loudspeaker systems like this are also used for distributed line systems. For example, they may be installed along the wall of an arena concourse, or outdoors, along a walkway leading to a theme park attraction.

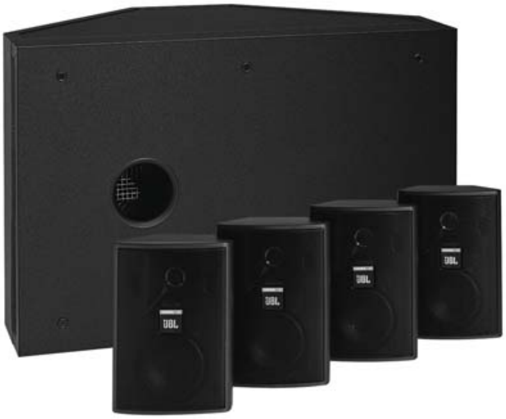


Figure 34-29. A family of packaged loudspeaker systems for distributed applications. Courtesy JBL Professional.

34.3.3.7 Designing the Distributed Ceiling System

Skipping Question 2 momentarily, here are discussions of the other three questions for a distributed ceiling system.

34.3.3.7.1 Question 1: Is It Loud Enough?

One of the advantages of a distributed system is that a typical listener is about the same distance from the nearest loudspeaker as any other listener. In addition, these distances are usually short compared to the critical distance D_c. Thus, the direct sound, not the reverberant sound, is the primary component of the L_P reaching the listener and the electrical power required (EPR) equation for the simplified system, Eq. 34-6, can be used. Use a single loudspeaker for this calculation for the minimum-overlap configurations (see Question 2, following). For the 50% overlap configurations, subtract 3 dB from the desired sound pressure level for the calculation of EPR to a single loudspeaker, since the equivalent of at least two loudspeakers will be covering each listener. If the room is highly reverberant and/or the ceiling height is sufficient to make the reverberant field a significant component of the sound at the listener's ears, the indoor EPR equation (Eq. 34-16 or 34-17) may be used. To include the effect of the multiple sources, use a value of $Q = 3/N$, where N is the total number of distributed loudspeakers in the critical distance D_c calculation. The number 3 is a typical Q for a distributed coaxial loudspeaker (an actual Q should be used if available). The value of N should be divided by two if each listener can hear the direct sound from two nearby loudspeakers and so on. Using the simplified method of Eq. 34-6 will always provide a safe answer to Question 1 since it considers direct sound only.

34.3.3.7.2 Question 3: Can Everybody Understand?

The *Alcons* equation (Eq. 34-18) works well for a distributed system. For the value of N, divide the total number of distributed loudspeakers in the room by the number of loudspeakers producing direct sound to a listener. Thus, if each listener is in the direct field of two loudspeakers, use a value of N equal to one-half the total number of loudspeakers and so on. Use a value of Q equal to the actual Q of each individual distributed loudspeaker. A good estimate for a coaxial ceiling loudspeaker is $Q = 3$. As with a central cluster system, try to maintain a 15 dB SNR and keep distortion, hum, and so on at a minimum for best intelligibility. In addition, remember that the *Alcons* equation works best for rooms with reverberation times of at least 1.6 s. In rooms with a lower RT_{60}, intelligibility is affected primarily by signal to noise.

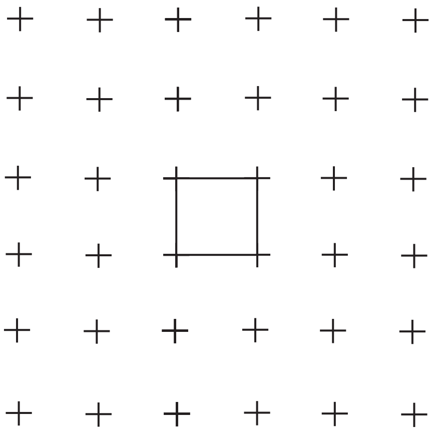
34.3.3.7.3 Question 4: Will It Feed Back?

Avoid placing microphones directly under a working loudspeaker. Provide switches to turn off loudspeakers above microphones when microphone positions will vary. Alternately, use an automatic mixer with logic outputs to automatically turn off loudspeakers above the active microphone. For conference rooms and other systems with fixed microphone positions, use an automatic microphone mixer with matrix output to create a mix-minus output signal routing system that always minimizes the signal from any microphone into a nearby loudspeaker (see Section 34.6.5). Use the PAG and NAG equations (Eqs. 34-19 and 34-20) with $Q = 3/N$ (see discussion under Question 1 previously) if the room has a significant reverberant component.

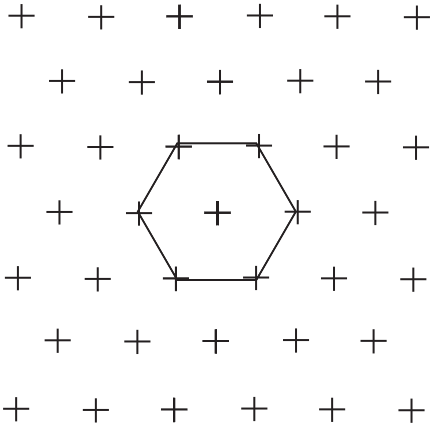
34.3.3.7.4 Question 2: Can Everybody Hear?

There are two basic patterns for laying out a distributed ceiling loudspeaker system. They are the square and hexagonal patterns, as shown in Fig. 34-30.

There are at least three variations of each of these two patterns, as shown in Figs. 34-31 and 34-32. The variations are in the spacing between the loudspeakers.



A. Square spacing.

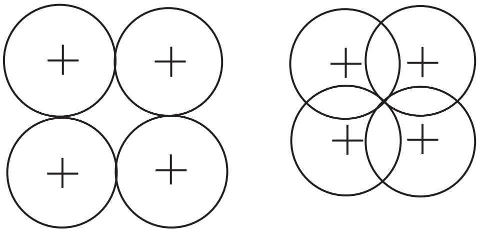


B. Hexagonal spacing.

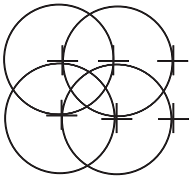
Figure 34-30. Square and hexagonal patterns for distributed loudspeaker systems. Courtesy Bosch/Electro-Voice.

An edge-to-edge spacing places the loudspeakers so that their coverage patterns just touch each other. A minimum-overlap spacing overlaps the coverage of the loudspeakers just enough to cover the dead spot in the edge-to-edge pattern. A 50% overlap is just that; each loudspeaker's coverage pattern overlaps the pattern of its neighbor by 50%. The result is that each loudspeaker is completely overlapped by a group of its neighbors.

The choice of one of these patterns should be made on the basis of the acoustics of the room, the ambient noise, and the type of listeners and talkers. In a difficult situation, such as might be encountered in a reverberant space with significant ambient noise and some listeners with hearing difficulties, a 50% overlap is indicated. For business music (background music) an edge-to-edge pattern may suffice.



A. Edge to edge. B. Minimum overlay.



C. Center to center (50% overlay).

Figure 34-31. Methods of square overlap for distributed loudspeaker systems.

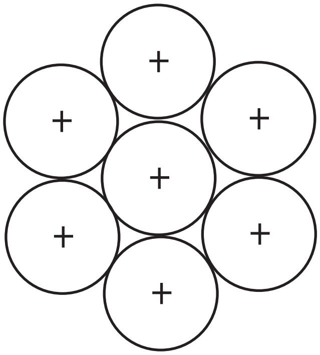
In any choice of coverage pattern, room obstacles, microphone locations, and seating area should be considered. There is no reason, for example, to cover wide aisles unless people will frequently be located there. In addition, remember that the coverage pattern should be calculated at about 4 ft above the floor for seated listeners or 5 ft above the floor for standing listeners.

34.3.3.7.5 Equalizing the Distributed Ceiling System

Equalization is discussed in more detail in Section 34.5.2.2. However, in general, the equalization process is the same as for a central cluster system. A typical listener position may be best chosen as in the overlap area of the loudspeakers for a 50% overlap system, or about 20° off-axis of a single loudspeaker for an edge-to-edge or minimum-overlap system. As in the central cluster process, choose several typical positions and equalize for a position that seems to be a good average as far as the before-equalization response. In an acoustically dry room (no significant reverberation field) the equalized response should show more high frequencies than the cluster system guidelines would indicate. This is because there is no low-frequency reverberation to boost the low frequencies artificially and bias the display on the real-time analyzer.

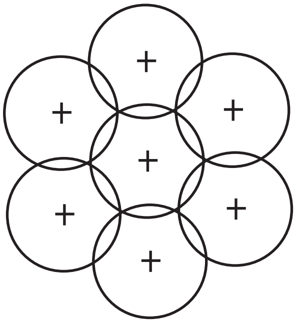
34.3.3.7.6 Distributed Systems in Rooms with Sloped Floors or Ceilings

The traditional approach for a system with sloped floors or ceilings is to divide the room into sections where the ceiling height is relatively constant and design a loudspeaker layout separately for each section. This will result in fewer loudspeakers per unit area in the higher-ceiling portions of the room as shown in Fig. 34-33. Additional power must be allocated to those loudspeakers in the high-ceiling portions.

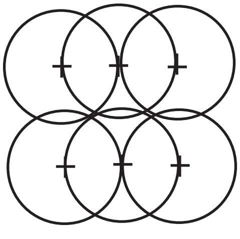


Square edge to edge	0.66	-3.69	4.35
Hexagonal edge to edge	0.95	-4.45	5.40
Square minimum	2.02	-0.02	2.04
Hexagonal minimum	1.36	-1.23	2.59
Square center to center	5.17	3.78	1.39
Hexagonal center to center	5.38	4.21	1.17

A. Edge to edge.



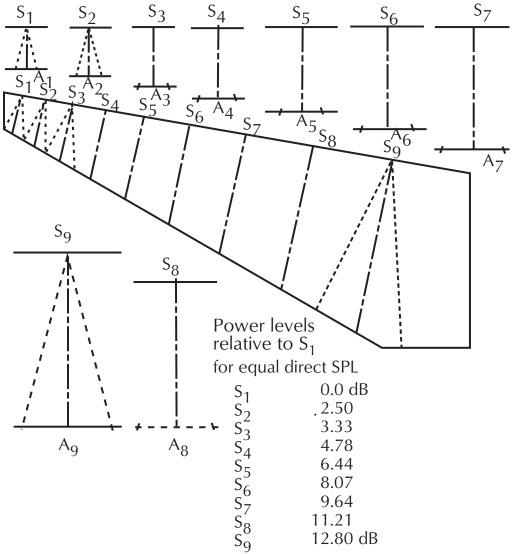
B. Minimum overlay.



C. Center to center (50% overlay).

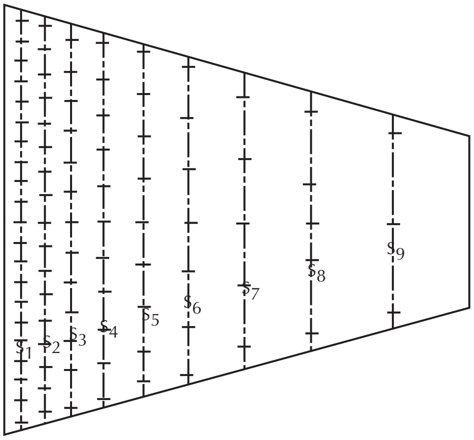
Figure 34-32. Methods of hexagonal overlap for distributed loudspeaker systems.

Another approach to a sloped-ceiling room is to place the loudspeakers as if the ceiling was flat at the lowest height and apply the same power to each loud-



S_1	0.0 dB
S_2	2.50
S_3	3.33
S_4	4.78
S_5	6.44
S_6	8.07
S_7	9.64
S_8	11.21
S_9	12.80 dB

A. Longitudinal section of room with loudspeaker row positions.



B. Plan view of room with loudspeaker positions.

Figure 34-33. Basic concepts of distributed placement in rooms with sloped floor and/or ceiling. Courtesy Bosch/Electro-Voice.

speaker. In the high-ceiling areas this approach covers each listener with additional loudspeakers. While the design is simplified and the layout more symmetric, this approach results in more loudspeakers, which will increase the cost.

34.3.3.8 Designing the Distributed Cluster System

A distributed cluster system consists of two or more clusters separated in space. Two clusters, for example, might be used to cover a long, narrow religious facility where a single cluster could not provide acceptable intelligibility in the rear seats. A second ring of clusters may also be installed in a large, fan-shaped room where the main system is an exploded cluster design. In this case, the second ring of clusters is installed on radii from the main clusters.

An examination of the *Alcons* equation (Eq. 34-18) shows that either increasing Q or decreasing D_2 will improve intelligibility. In the case where a loudspeaker with high enough Q will not provide wide enough coverage or where a loudspeaker with high enough Q is simply not available, adding one or more additional clusters that are closer to the far listeners may be the answer since this decreases D_2.

Design the first cluster (or exploded clusters) to cover the seating areas out to (and slightly beyond) the position of the second cluster (or second ring of clusters). All other design criteria for the first cluster remain the same as if it were the only cluster. The value of N (the D_c modifier) for either cluster must include the effects of both clusters, however. Design the second cluster to cover the remaining seating area to the farthest listener. In many systems of this type, the second cluster can have a reduced low-frequency section for frequencies below about 200 Hz. This is because the frequencies below 200 Hz do not contribute to intelligibility and because the reverberant field in most rooms requiring a second cluster will carry the low frequencies to the farthest listener with no need for reinforcement from the second cluster.

In calculating *Alcons* (or *EPR* or *PAG* and *NAG*) for either cluster, the value of N must take into account the loudspeakers in both clusters, although the value of D_2, of course, will be shorter than it would have been for a single cluster in the same room.

34.3.3.8.1 Signal Delay in a Distributed Cluster System

The second cluster in the previous example must receive a signal that is electronically delayed from the signal sent to the first cluster. See Section 34.3.3.8.1 for an example of calculating this delay.

34.3.3.8.2 Distributed Clusters with No Delay

In a circular stadium or on long, narrow bleachers such as at a race track, a system of distributed clusters may provide the best coverage and may not require electronic delay because the sound reaching each listener is primarily from one cluster and any nearby cluster is essentially the same distance from the listener (and thus not acoustically delayed). Design of each cluster in such a system is straightforward. Choose a seating area that can be easily covered by a single cluster. Calculate the difference in distance from a typical listener to adjacent clusters. Avoid wide spacing between clusters that could cause a listener hearing two clusters to hear the second cluster as an echo of the first. Provide sufficient overlap between coverage areas to insure adequate sound pressure level to all listeners but avoid wide overlap areas to prevent the problems inherent in the split cluster discussed in Section 34.3.2.3.

This same approach applies to smaller loudspeakers distributed around the concourse in an arena or along the sidewalk leading to a theme park attraction.

34.3.3.8.3 Equalizing the Distributed Cluster System

Equalization is discussed in more detail in Section 34.5.2.2. However, if all clusters in a distributed cluster system are the same and are covering areas with similar acoustics, equalization may be performed for a single cluster and duplicated for the other clusters. Check the response of the other areas to confirm the equalized curve is similar. If clusters are covering acoustically different areas or if they are designed for different loudspeakers, each type of area or loudspeaker must receive separate equalization (the central cluster plus under-balcony distributed system, for example).

34.3.4 Crossover Networks and Biamplification

Loudspeaker crossover networks are also discussed in Chapter 17.

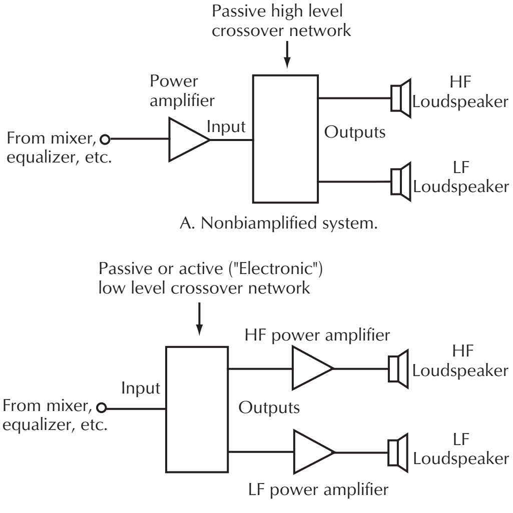
34.3.4.1 Definitions

Crossover Network. A *crossover network* is a filter network that routes high frequencies to a high-frequency loudspeaker and low frequencies to a low-frequency loudspeaker. If the crossover network is part of a biamplified system, it will do its frequency division prior to the power amplifiers. Three-way and four-way crossovers perform the same function but divide the frequencies into more sections.

Passive Device. A *passive device* uses no active components (tubes, transistors, ICs) and needs no power supply (ac, dc, battery). The crossover network in a typical packaged loudspeaker system is a passive device.

Active Device. An *active device* uses one or more active components and requires some type of power supply. An electronic crossover, used in a biamplified system, is an active device.

Biamplified System. A *biamplified system* uses an electronic crossover (commonly a module in a DSP) and it uses separate power amplifiers for the high- and low-frequency loudspeakers, Fig. 34-34. A *triamplified system* is a three-way loudspeaker system with a three-way electronic crossover and separate power amplifiers for the low-, mid-, and high-frequency loudspeakers. To simplify, we often speak of triamplified and multiamplified systems as being *biamplified*.



A. Nonbiamplified system.

B. Biamplified.

Figure 34-34. Biamplified and nonbiamplified systems.

Head Room. *Headroom* is the difference, in decibels, between the peak and rms levels in the program material.

34.3.4.2 Advantages of a Biamplified System

One advantage of a biamplified system is that it can actually provide more head room per watt of amplifier power than a system with a traditional (loudspeaker level) passive crossover.

The reason this happens is that most music, especially popular music, is bass heavy; that is, there is much more energy at low frequencies than at high frequencies. When both high and low frequencies are present in a program, the high-energy bass frequencies will dominate the output of the system power amplifier leaving little or no power for the high frequencies. The result can be severe amplifier clipping (distortion) of the high-frequency material. By biamping the system, with an electronic crossover, the high-frequency material can be routed to its own power amplifier avoiding the clipping problem. This results in an effective increase in head room that is greater than that which would be obtained by simply using a single power amplifier of equal power output.

Another advantage of biamplification is that it does not absorb amplifier power as a loudspeaker-level passive crossover does. Biamplification, by removing this loudspeaker-level crossover, improves the overall system efficiency.

Improved damping factor is another advantage of biamplification. The damping factor of a power amplifier is a number found by dividing its load impedance (the impedance of the loudspeakers) by the actual output impedance of the amplifier, which will be very low for a modern solid state power amplifier. An amplifier with a high damping factor can exert a greater control over the motions of a loudspeaker cone than an amplifier with a low damping factor. Thus, a high damping factor may improve the sound quality of a system. A loudspeaker-level passive crossover lowers the damping factor by inserting its impedance between the amplifier and the loudspeakers. Removing the loudspeaker-level passive crossover, and biamplifying the system, can thus improve the damping factor.

Biamplification can lower distortion by increasing head room as explained previously. However, if clipping distortion occurs anyway, it may be less audible in a biamplified system. In a conventional, nonbiamplified system, the high-frequency harmonics generated by clipping of a low-frequency transient are passed through the loudspeaker-level crossover to the high-frequency loudspeaker where they will be quite audible. In a biamplified system, there is no crossover and no high-frequency loudspeaker after the low-frequency power amplifier. Thus, the clipped low-frequency signals and their harmonics are restricted to the low-frequency loudspeaker and, due to its poor high-frequency response, the low-frequency loudspeaker attenuates the audibility of these unwanted harmonics.

34.3.4.3 When to Use a Loudspeaker System with Passive Crossover

In small sound systems where high sound levels aren't needed and economy is a major consideration, a loudspeaker system with a traditional passive crossover network may be the best choice. In addition, the crossover in a packaged loudspeaker system from a manufacturer usually includes a certain amount of equalization designed to improve the overall response of the loudspeakers. Biamplifying this system would require the addition of a graphic or parametric equalizer and, perhaps, a signal delay in addition to the additional power amplifier and electronic crossover. Thus, it's usually best to use the manufacturer's loudspeaker-level passive crossover in a packaged system unless the manufacturer has made specific provision for biamplification and offers recommended DSP settings for biamplified operation.

34.3.4.4 Signal Alignment in a Loudspeaker Crossover Network

In the crossover frequency range the output from the loudspeaker system includes output from both the high- and low-frequency components. If the arrival time, at the listener's ears, of the signal from the high-frequency component differs from the arrival time of the signal from the low-frequency component, significant frequency response degradation can result near the crossover frequency.

The solution to this problem is to physically or electronically align the low- and high-frequency components so that their signal arrivals coincide at the listener's ears.

34.3.5 Protecting the Loudspeakers

34.3.5.1 Loudspeaker Failure Modes

The discussions in this section apply equally to both low-frequency cone-type loudspeakers and high-frequency compression drivers.

Discounting manufacturing defects that may cause random failures, loudspeakers normally fail from either excessive average power or from excessive peak power at low frequencies. Loudspeakers may also fail due to materials aging, physical damage, weather-related deterioration or damage from insects or other pests.

Excessive average power causes voice coil heating and eventually voice coil failure (or failure of other components in the voice coil area). Excessive low-frequency peak power causes mechanical failure of the loudspeaker due to overexcursion. The voice coil may separate from the rest of the loudspeaker or the loudspeaker cone (or driver diaphragm) may tear or shatter. Protecting a loudspeaker, then, is primarily a matter of preventing these two failure modes and protecting it from weather, physical damage, and pests.

34.3.5.1.1 Choosing Power Amplifiers to Prevent Excessive Average Power

It is possible to destroy a loudspeaker by using a power amplifier that is too large (one whose power output exceeds the power capacity of the loudspeaker by some margin). It is also possible, under some conditions, to destroy a loudspeaker by using a too-small amplifier.

The too-small amplifier is one that does not have enough power output to meet the requirements of both the needed L_P at the farthest listener and the system head room (from the electrical power required equation). Attempting to reach the system requirements will exceed output capabilities of the amplifier, which will cause the amplifier to clip, turning sine waves into semisquare waves and vastly increasing distortion levels.

This clipping causes two problems. First, the square wave can actually draw twice the power output from the power amplifier. That is, if the amplifier is rated at 100 W, a full-voltage square wave can cause the amplifier to deliver as much as 200 W, depending on power-supply limitations and the internal protection circuits of the amplifier. This double power output can be a threat to the loudspeaker all by itself. Second, the square wave causes the loudspeaker cone/diaphragm to move outward (or inward) and stay there for awhile, then move in the other direction and stay there for awhile. When the loudspeaker cone/diaphragm is not moving (at the top and bottom of the square wave), the energy supplied to the voice coil is being entirely converted into heat, with obvious consequences.

Thus, one way to prevent loudspeaker damage is to use a power amplifier that has enough output to reach the maximum L_P requirement and the system head room requirement. Fortunately, in most sound reinforcement systems, actual electrical power required is small, and therefore the problem of a too-small power amplifier shows up primarily in large sound reinforcement systems or in popular music-oriented entertainment systems.

Loudspeaker power capacity is usually rated using some type of noise with a specified head room factor. For example, a loudspeaker may be rated at 100 W continuous pink-noise, with a 10 dB crest factor from 50–1000 Hz. That crest factor is the difference between the average and peak power in the pink-noise signal. The crest factor concept is similar to head room in the sound reinforcement system. Thus, in theory, this loudspeaker can be fed 100 W of pink-noise, band limited from 50–1000 Hz, and with peaks that reach 1000 W. The 100 W loudspeaker is, theoretically, safe with a 1000 W amplifier if system head room is kept at 10 dB or above. In practice, of course, it is very risky to power a 100 W loudspeaker from a 1000 W amplifier. The reasons are many but include the operator who will push the system past its design limits, ignoring distortion and the possibility of sustained feedback that can draw full power from the power amplifier for an extended period.

What then is a safe power amplifier size for a 100 W loudspeaker? In most systems, about twice the rated power capacity of the loudspeaker will be safe provided other potential problems are considered, as discussed in Section 34.3.5.1. In addition, the amplifier should, as previously discussed, be at least capable of supplying enough power to meet the system maximum L_P and head room needs. Full-power, sustained feedback can still, of course, destroy the loudspeaker, but at only twice the rated power capacity of the loudspeaker, the system will likely sound very distorted before the loudspeaker is in danger. This should prompt the system operator to turn down the level, preventing damage.

34.3.5.1.2 Loudspeaker Power Capacity Specifications

Besides pink-noise power capacity, manufacturers commonly use several other power capacity rating methods. Variations on the concept of program power are used in an attempt to define the loudspeaker's power capacity when the source is normal program material. The interpretation of normal program material, of course, depends on the manufacturer and it is common for this power capacity rating to be significantly higher than other ratings.

Rms power is another common rating method. Mathematically, there is no such thing as rms power. rms power is calculated from rms voltage and load resistance. The correct term should be *average power*. However, an rms power rating for a loudspeaker is usually similar to a pink-noise rating.

The EIA (Electronic Industries Association) has a loudspeaker power capacity standard using shaped noise that is similar to a pink-noise rating with low- and high-frequency roll-off. This standard, known as EIA RS-426A, is a reasonably reliable indication of the loudspeaker's thermal power capacity and is similar to a pink-noise power capacity.

There are other methods of rating loudspeaker power capacity (see Chapter 17), but because of the complexity of the subject, all require some interpretation. It may be that a single number is simply not sufficient to fully describe a loudspeaker's power capacity.

34.3.5.2 Protecting against Excessive Low-Frequency Peak Power

The cone/diaphragm excursion of a loudspeaker increases at low frequencies. The exact amount of increase depends partly on the enclosure or horn the loudspeaker (or high-frequency driver) is used with. Nevertheless, there is some frequency below which each loudspeaker/enclosure or driver/horn should not be used. This low-frequency limit is normally given in the manufacturer's specifications or, for a vented enclosure design, may be estimated as f_b (the vented box resonant frequency). At very low input power, frequencies lower than this limit will not cause damage. At normal to high power inputs, however, low frequencies can cause loudspeaker damage due to overexcursion.

The cure for this overexcursion is simply to prevent these low frequencies from ever reaching the loudspeaker by using some type of high-pass filter. This may be in the form of a system crossover network, which prevents low frequencies from reaching the high-frequency loudspeaker, or in the form of a separate high-pass filter (often part of a graphic equalizer or DSP), which prevents very low frequencies from reaching the low-frequency loudspeaker. One valuable protection device is a series capacitor, used on the high-frequency loudspeaker, which can reduce the effects of any low frequencies that may pass through the power amplifier due to such problems as turnon/turnoff transients.

Significantly, excessive power input to a loudspeaker at frequencies above its rated frequency range can also be dangerous. Since the loudspeaker cannot produce sound from these frequencies, the input power is mostly converted into heat, adding to the potential problem of excessive average power.

34.3.5.3 Loudspeaker Protection Devices

Careful system design and operation by an experienced operator are the best protection against loudspeaker failure. The following devices can help, however, and may be used in almost any system design.

34.3.5.3.1 Fuses

Fuses are poor loudspeaker protection devices. Standard fuses may be capable of protecting a loudspeaker against excessive average power, but they are too slow to protect a loudspeaker successfully against sudden peaks. Fast-blow instrumentation fuses, with improved time response, may blow on normal program peaks and needlessly disrupt sound system operation. Slow-blowing fuses, on the other hand, may not blow quickly enough to prevent loudspeaker damage due to voice coil overheating.

Despite these limitations, fuses are sometimes used as loudspeaker protection devices. If fuses are used, fuse each loudspeaker separately so that a single fuse failure will not completely interrupt system operation. Choose a starting fuse size from the following equation:

$$F = 0.75\sqrt{\frac{P}{Z}} \tag{34-29}$$

where,

F is the fuse size in amperes,

P is the rated power capacity of the loudspeaker,

Z is the rated impedance of the loudspeaker.

This equation gives a fuse size that will blow when the input power to the loudspeaker reaches 75% of its rated value. Fuse size may be increased if this fuse blows frequently, but avoid fuses larger than about twice this value since they will pass enough current to overpower the loudspeaker.

Early direct-coupled power amplifiers, when they failed, would often pass their full dc supply voltage to the loudspeaker. This voltage can result in loudspeaker/driver voice coils that are described as being "french fried." A fuse will help protect a loudspeaker against this type of power amplifier failure mode. On high-frequency loudspeakers, however, a capacitor is probably a better protection device against this problem. In addition, most modern power amplifiers have some kind of internal protection (such as an output relay) that should prevent the problem of dc at the output even when the amplifier itself fails.

34.3.5.3.2 Capacitors

A series capacitor (connected electrically in series with the loudspeaker's positive input lead) can help prevent excessive low-frequency power and can protect the loudspeaker against dc power from a faulty power amplifier. Capacitors can be chosen from

$$C = \frac{500{,}000}{\pi f Z} \tag{34-30}$$

where,

C is the value of the capacitor in microfarads,

Z is the rated impedance of the loudspeaker,

f is a frequency chosen as follows:

- If the system is two way (or three way, and so on), choose f equal to one-half the system crossover frequency. If the capacitor is to be used as a high-pass protection device in a voice-only system, choose f equal to the desired high-pass frequency, remembering that a single series capacitor provides about a 6 dB/octave slope rate.
- Choosing a low-frequency value for f in the equation results in a very large capacitor. Thus, a capacitor is usually an impractical method of protecting a low-frequency loudspeaker. A high-pass filter prior to the power amplifier is better for a low-frequency loudspeaker.
- For a high-frequency driver, using the driver's rated impedance for Z may result in errors since the actual impedance can be much higher at frequencies below crossover. Thus, it may be a good idea to actually measure the driver's impedance at the chosen frequency f and use this in the equation.
- Choose a capacitor with a voltage rating at least equal to the maximum peak-to-peak voltage output of the power amplifier. This will be the sum of the absolute values of the positive and negative power supply voltages for a direct-coupled amplifier and can be approximated from the following equation for either direct-coupled or transformer-coupled amplifiers:

$$V_{P-P} = 2.828\sqrt{PZ} \tag{34-31}$$

where,

V_{P-P} is the peak-to-peak voltage output of the power amplifier,

P is the rated output power of the amplifier,

Z is the rated load impedance of the amplifier,

the value of 2.828 is twice the square root of two.

The value of V resulting from this equation may be conservative, since most amplifiers can produce power output in excess of their rated value for short periods. Thus, the actual voltage rating for the capacitor should probably be somewhat higher.

The capacitor must be nonpolarized. Motor run types are considered good choices for sound system applications. Motor start capacitors may be used. Standard electrolytic capacitors, if nonpolarized, can be used, but these capacitors normally have a very poor tolerance in actual capacitance value and, thus, may not provide the expected protection.

34.3.5.3.3 Limiters

A limiter is not normally considered a loudspeaker protection device, but it may be one of the best and most practical. The limiter, Fig. 34-35, can be adjusted to prevent the system power amplifier from exceeding its power output capabilities and can help prevent high-power peaks from reaching the loudspeakers. In systems where sound quality is a primary consideration, adjust the limiter so that its threshold is high and its compression ratio is high. This way, the limiter will not be in operation until a potentially dangerous peak is detected. Then, the high compression ratio of the limiter will clamp the peak and help prevent loudspeaker damage.

34.3.5.3.4 Other Protection Devices

Transformers can help protect loudspeakers because they cannot pass dc. In the past, some transformers included series capacitors to limit the low-frequency energy to a high-frequency driver. Autotransformers, on the other hand, may pass dc to a loudspeaker since they have only a single winding.

Passive crossover networks, because they include one or more series capacitors, provide good protection for high-frequency drivers and some protection for the low-frequency loudspeaker (against excessive high-frequency power levels). The passive crossover networks used in some packaged loudspeaker systems include sophisticated protection circuitry. The manufacturer will normally specify this in its sales literature and intruction manuals.

High-pass and low-pass filters, similar to those often found in a DSP or on a graphic equalizer, are valuable in any system. A high-pass filter helps keep out unwanted low frequencies that could cause overexcursion. A general rule is that, except for subwoofer systems, a 40–160 Hz high-pass filter should be used in all systems. Even for subwoofers, a 10–20 Hz (or higher) high-pass filter can help prevent dangerous overexcursion. High-pass filters are often available on mixer input channels. Using them here can help reduce damage from dropped microphones or other problems. Low-pass filters help prevent heat-producing radio-frequency energy (picked up from outside sources or from faulty system electronics) from reaching the loudspeakers. Low-pass filters also keep out audio frequencies above the loudspeaker's range (which would also cause unwanted heating).

dbx
160A
BYPASS
SLAVE
THRESHOLD
BELOW
OVEREASY
ABOVE
COMPRESSION RATIO
OUTPUT GAIN
INPUT
OUTPUT
LEVEL
GAIN REDUCTION
DISPLAY
COMPRESSOR/LIMITER

Figure 34-35. A versatile high-quality compressor/limiter. Courtesy dbx Professional.

Special-purpose DSP processors included with some packaged loudspeaker systems often include sophisticated loudspeaker protection including limiters and even sliding high and low-pass filters.

34.3.5.3.5 Protecting against Weather, Physical Damage and Pests

Physical damage may be caused by overexcited fans at a sporting event or by vandalism. When possible, locate the loudspeakers out of the reach of potential vandals. In some facilities, it may be necessary to build protective cages to prevent damage or theft. Some manufacturers offer vandal-resistant loudspeakers for use in correctional facilities and schools.

Weather and pest damage can best be avoided by choosing loudspeakers designed to resist these problems. When possible, locate loudspeakers in protected areas such as under a balcony or awning. Loudspeakers in outdoor summer amplitheaters may be removed and stored for the winter season or covered for protection from winter damage.

34.3.5.3.6 Age-Related Loudspeaker Damage

The foam surrounds on some cone-type loudspeakers will deteriorate and fail after 10–15 years of use. It's best to choose loudspeakers with long-lasting surround materials, such as impregnated cloth, to avoid this problem. It's also possible for cones to age and sag after many years, causing the voice coil to rub against the pole piece. When this happens, recone or replace the loudspeaker.

34.4 Electronic Components for Sound Reinforcement

34.4.1 General Specifications fozr Sound Reinforcement

Electronic devices for professional and commercial sound reinforcement systems should have balanced inputs and outputs. Line level outputs should be +4 dBu nominal with 20 dB head room for a +24 dBu peak output level. Many digital audio devices have a maximum output of +18 dBu, which is acceptable if the system is designed for this maximum level. The devices should be rack mountable except for those intended to mount on desklike mixing consoles. They should utilize high-quality electronics with low levels of hum and noise; wide smooth frequency response; and low distortion. For most applications, the devices should conform to a recognized safety listing such as UL (Underwriter's Laboratories).

34.4.2 Mixers and Mixing Consoles

See Chapter 25 for a thorough discussion of all kinds of mixing consoles. There are several types of mixers and mixing consoles commonly used in sound reinforcement systems. Simple rack mixers, like the one shown in Fig. 34-36, may be all that's needed for a college lecture hall or for a religious facility with a spoken worship style. Automatic mixers, as discussed in Section 34.4.3, can reduce the need for an operator in these simple systems.

Choose a desk-type mixing console, like the one in Fig. 34-37, for live theater or for any facility that hosts live entertainment events. Modern religious facilities often include dramatic and musical performances as part of their worship services. These facilities need a desk-type mixing console. The versatility of a desk-type mixing console means the operators must be well trained in the artistic and technical aspects of its operation.

Tour sound systems use desk-type mixing consoles. Larger tour sound systems may also use a special-purpose type of mixing console known as a monitor mixer. A monitor mixer is specifically designed to mix the monitor loudspeakers on a performance stage. For this reason, the stage monitor mixer is normally located at stage left or stage right where the operator can see and hear the monitor loudspeakers.

Digital mixing consoles, like the one shown in Fig. 34-38, perform the same functions as their analog cousins but they have additional features, such as memory scenes and multifunction outputs, that cannot be implemented on analog consoles. A memory scene stores most or all settings on the console in a digital memory location for recall at the touch of a button. A

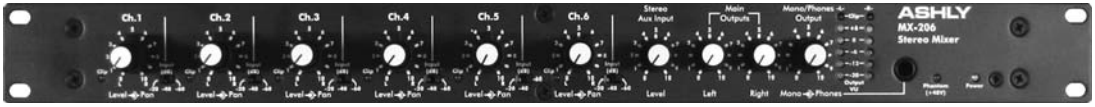


Figure 34-36. Simple rack mixer. Courtesy Ashley Audio.

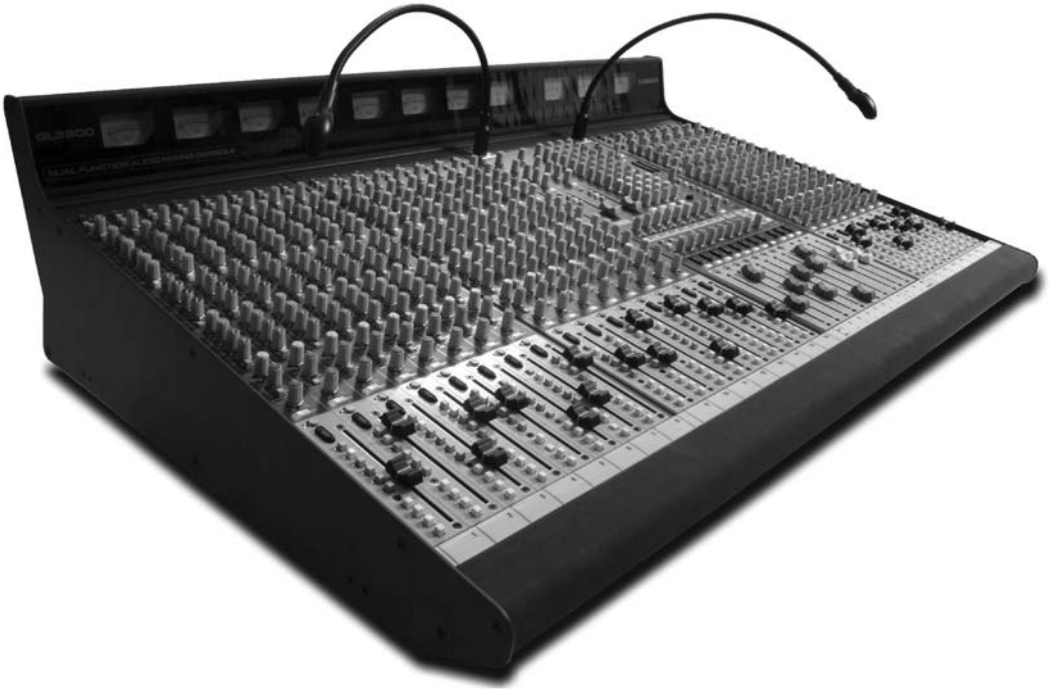


Figure 34-37. A sound reinforcement mixing console. Courtesy Allen and Heath.

live theater can use this to set up console settings for each scene during rehersal and call them up quickly during a performance. A church with traditional and modern services can set up the console for each service and switch between them quickly. This can be a great help when operators are inexperienced. Multifunction outputs can work as group, main, matrix, or monitor outputs depending on configuration. There are many other useful features on digital mixing consoles. However, some digital mixing consoles are so feature-rich that they are confusing to an inexperienced operator. Control functions may change when memory scenes are changed and, unlike an analog mixing console, it's generally not possible to understand the configuration and settings of the console by simply looking at its controls. Thus, digital mixing consoles are best suited to facilities with experienced operators, Fig. 34-38.

34.4.2.1 Mix Groups, Auxiliary Groups and Matrix Mixing

Most mid- to large-mixing consoles have mix groups and auxiliary groups. Larger mixing consoles often have an output matrix mixing section. Experienced operators develop a mixing style that uses all of these features effectively. Here is one common approach, Fig. 34-39.

Start by connecting sources (microphones, musical instruments, etc.) to the inputs in a way that makes it easy to reach the most-used controls. Set up the input gains and losses for minimum noise and maximum head room (see Chapter 28). Then assign similar sources to the various mix groups in a logical manner. In a religious facility, for example, the operator could assign spoken voices to group 1, singing voices to group 2, choir to group 3, instruments to group 4, percussion to group 5, electronic organ to group 6, and so on. This allows the operator to raise or lower the volume of each

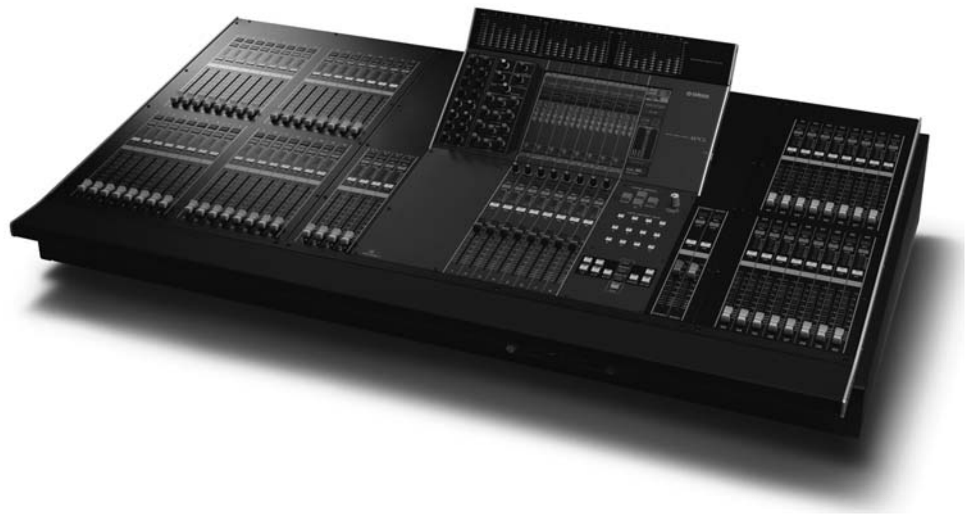


Figure 34-38. A digital mixing console. Courtesy Yamaha.

group with a single fader. Assign the groups to the master left and right outputs to feed the auditorium loudspeaker systems.

For individual channel special effects, use that channel's insert feature. For special effects on a group of sources, use an auxiliary group. In the religious facility example, certain spoken voices may benefit from artificial reverberation during a dramatic presentation. Assign these voices to an auxiliary group and feed the aux group output to the reverberation device. Return the output of the reverberation device to an unused input channel or other available input.

For stage monitor mixing, assign selected inputs to an auxiliary group to feed a stage monitor loudspeaker. By using two or more auxiliary groups, the operator can provide customized mixes for different needs on the stage. In the religious facility example, the choir needs to hear the spoken voices and musical instruments, but may not need to hear itself in the monitor mix. In contrast, the pastor or a lay reader needs to hear the choir and musical instruments but does not need to hear the spoken voices.

If the mixing console has a matrix output, use this section to feed the various loudspeaker systems. Assign groups (and aux groups) to the matrix outputs to achieve an optimum mix for each loudspeaker system. In the religious facility example, use matrix outputs 1, 2, and 3 to feed the auditorium left, center, and right clusters. Use matrix output 4 to feed the under-balcony loudspeakers. Use matrix output 5 to feed any external overflow rooms, mothers' rooms, and offices. Use matrix outputs 6 and 7 to feed a stereo recording or live broadcast feed.

By using the matrix outputs in this manner, each loudspeaker system or recording or broadcast feed can have a custom mix. In the religious facility example, the auditorium loudspeaker clusters need all of the groups except the electronic organ, which has its own loudspeaker system in the auditorium. The recording or broadcast feeds, and the overflow room feeds, however, need the organ and perhaps an audience response microphone feed. Custom mixes like this can be set up in the matrix and need very little adjustment during a performance.

34.4.3 Automatic Microphone Mixing

Many of the tasks a human operator performs on a simple mixer are predictable. For example, the human operator turns up the volume controls for microphones that are in use and turns down the volume controls for microphones that are not in use. In addition, an experienced human operator will turn down the master volume control about 3 dB each time the number of in-use microphones doubles to help avoid feedback from the NOM problem discussed in Section 34.2.2.5. An automatic mixer performs these two functions without the aid of an operator, Fig. 34-40.

The first commercially successful automatic mixer was invented and patented by Dan Dugan, a consultant in San Francisco, and marketed by Altec Lansing Corp. The Dugan automatic mixer exclusively used analog

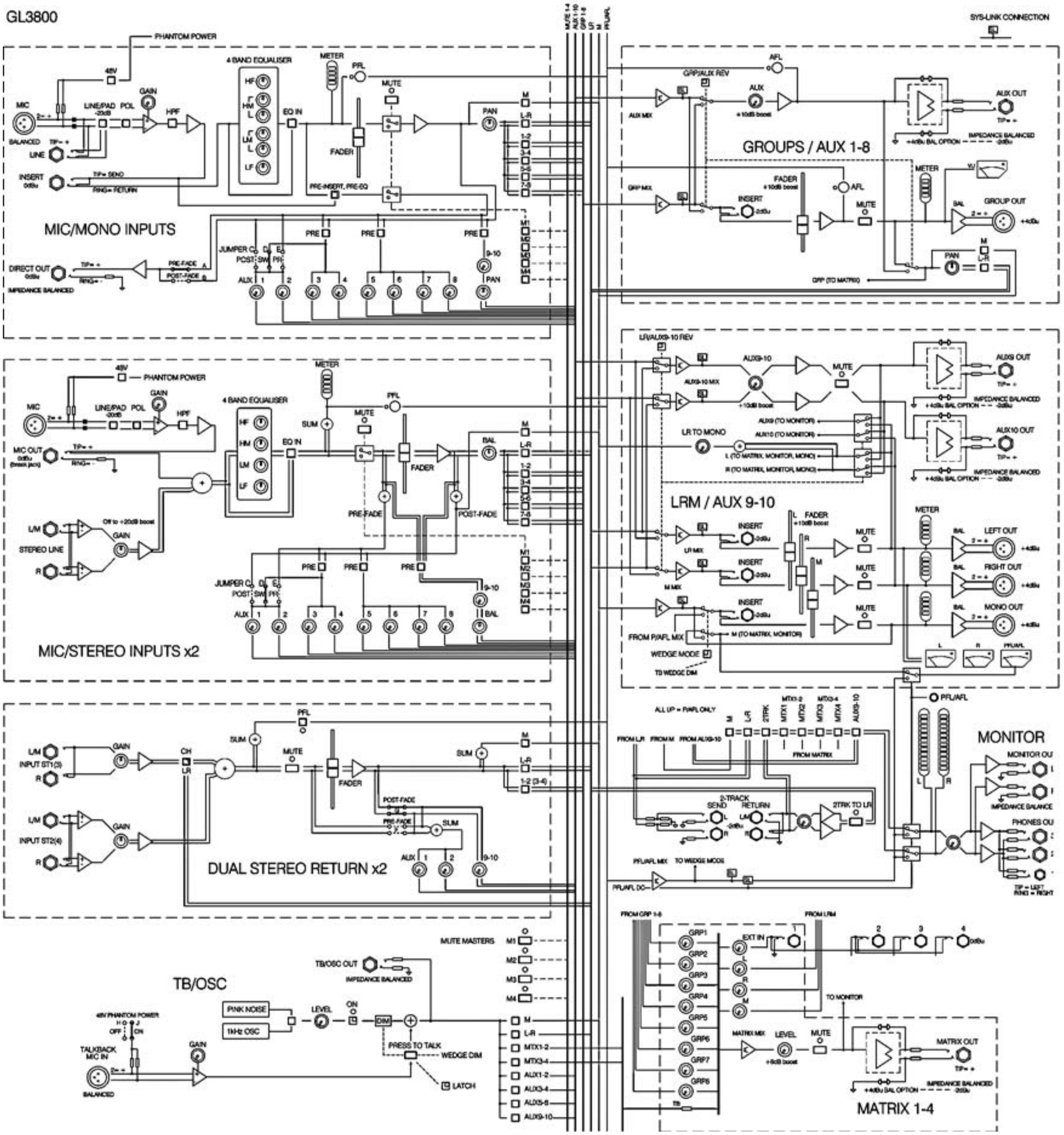

GL3800
SYS-LINK CONNECTION
PHANTOM POWER
48V
MIC
LINE/PAD -20dB
POL
GAIN
HPF
4 BAND EQUALISER
HF
HM
LM
LF
EQ IN
METER
PFL
MUTE
FADER
PAN
INSERT
LINE
DIRECT OUT
IMPEDANCE BALANCED
MIC/MONO INPUTS
PRE
AUX
GROUPS / AUX 1-8
AUX OUT
GROUP OUT
GRP (TO MATRIX)
MIC/STEREO INPUTS x2
MIC OUT
STEREO LINE
PRE-FADE
POST-FADE
BAL
LRM / AUX 9-10
AUX9 OUT
AUX10 OUT
LR TO MONO
LEFT OUT
RIGHT OUT
MONO OUT
FROM PFL/AFL MIX
WEDGE MODE
TO WEDGE DIM
INPUT ST1(3)
INPUT ST2(4)
DUAL STEREO RETURN x2
SUM
ALL UP = PFL/AFL ONLY
FROM MATRIX
2-TRACK SEND
RETURN
2TRK TO LR
PFL/AFL
MONITOR
MONITOR OUT
PHONES OUT
MUTE MASTERS
TB/OSC
TB/OSC OUT
PINK NOISE
1kHz OSC
LEVEL
ON
DIM
PRESS TO TALK
WEDGE DIM
LATCH
TALKBACK MIC IN
48V PHANTOM POWER
FROM GRP 1-8
FROM LRM
EXT IN
TO MONITOR
MATRIX MIX
MATRIX OUT
MATRIX 1-4

Figure 34-39. GL3800 block diagram of a mixing console with matrix outputs. Courtesy Allen and Heath.

circuitry to perform the automatic mixing process according to

$$L_{N'} = L_N - [Sum(L_N) - L_N] \qquad (34\text{-}32)$$

where,

L_N is the level in an individual mixer channel before the automatic circuitry has attenuated that channel,

$L_{N'}$ is the level in the channel after attenuation,

$Sum(L_N)$ is the sum of the levels in all channels before they are attenuated,

all values are in decibel notation.

In effect, the equation says that each individual input channel is attenuated by an amount in decibels equal to the difference in decibels between the level of that channel and the sum of all channel levels.

The significance of this equation is that, while it performs the two functions mentioned previously, it does not mention the word threshold nor the word switch. A Dugan system automatic mixer varies the microphone levels in a continuous manner depending only on the relationship of each individual channel level to the sum of all the channel levels.

A user sets up the Dugan mixer by adjusting each individual volume control to a position suitable for the person talking. That means the volume control for a quiet talker's microphone will have a higher setting than that for a loud talker. These volume control settings assure that the circuitry treats all microphones equally in the equation. After this initial volume control adjustment, the user ceases interacting with the mixer. New talkers, of course, or significant changes in talker input level, require human intervention. A dummy microphone modification helps keep the mixer's automatic circuitry from being fooled by ambient noise.

Other automatic mixers switch microphones on (to a volume control level preset by the user) when someone talks into the microphone and off when no one talks into the microphone. They also reduce the master volume control by approximately 3 dB when the number of in use microphones doubles. Most of these mixers incorporate sophisticated digital circuitry to make the decisions about when to turn a microphone on or off and exactly how much to attenuate the master volume control. As a result, a well-designed automatic mixer of either type can be successfully used in a system designed for automatic mixing.

Today, automatic mixers may be digital and automatic mixing functions are often included in multichannel DSP devices.

34.4.3.1 Special Features

A number of additional functions/features are either standard or optional on most automatic mixers. Specific features, of course, depend on the make and model chosen.

For example, on the switching-type mixers, users may have the option of adjusting the threshold setting. The threshold is the L_P at which the mixer turns on a microphone channel. In very high noise areas, for example, a user could increase the threshold to reduce the problem of microphones turning on from ambient noise input. Another feature available on some mixers is adjustment of the amount of off attenuation. That is, the off state can be redefined from no attenuation at all to infinite attenuation (true off). By selecting no attenuation, the on-off switching feature is defeated, and the mixer functions only as a number of open microphones (NOM) attenuator.

One valuable feature found on most mixers is a logic output. This logic output is a dc voltage output, usually compatible with TTL circuitry levels, and it goes high when the microphone is on and goes low when the microphone is off. This logic output can be used to activate relays for zone paging or to activate complex microphone priority switching in conference systems.

Most automatic mixers also allow the user to defeat the automatic circuitry on an individual input channel. This allows a tape machine or other nonmicrophone input to be added to the mix without affecting (and without being affected by) the automatic mixing of the system microphones.

34.4.3.2 Applications for Automatic Mixing

In systems with undemanding, predictable mixing requirements, the automatic mixer may be able to completely replace the human operator. Examples are conference and courtroom systems and speech-oriented systems in religious facilities. In these systems, the installer sets up the system volume controls and instructs the user simply to turn the entire system on and off since the automatic mixer will take care of everything else.

In actuality, systems like these are rare. More common are systems where the automatic mixer becomes an operator aid rather than completely replacing the human mixer. Any of the previously mentioned systems where different talkers use the same microphone require some human intervention. But the automatic mixer can also aid the human operator in more sophisticated systems including entertainment-oriented systems and dramatic (live) theater presentations.

Most automatic mixers are unsuitable for mixing musical material. However, most can be used effectively for voice mixing of footlight microphones in a theater, and some automatic mixers may find use in submixing of instruments or vocals in an entertainment-oriented system. In all cases, the ability of the mixer to sense in-use microphones and attenuate (or to turn off) other microphones is a valuable aid in reducing unwanted noise pickup. In addition, the ability to help reduce the possibility of feedback (the *NOM* function) is welcome in any system.

34.4.3.3 Mix-Minus and Matrix Mixing

The automatic mixer in Fig. 34-40 includes a matrix output. This enables a mix-minus system for conference rooms or other multimicrophone systems, Fig. 34-41. Each talker at this conference table has his or her own microphone and loudspeaker. The signal from a given talker's microphone is amplified to a greater degree in loudspeakers that are farther away from the talker. This is a natural way to make the talker's voice heard well at any point around the conference table. In addition, this system helps control feedback because the talker's voice is not amplified into his or her local loudspeaker and only slightly into nearby loudspeakers. In combination with normal automatic mixer functions, the mix-minus approach can make effective sound reinforcement possible in a large conference room. This approach is also valuable for audio or video teleconferencing systems.

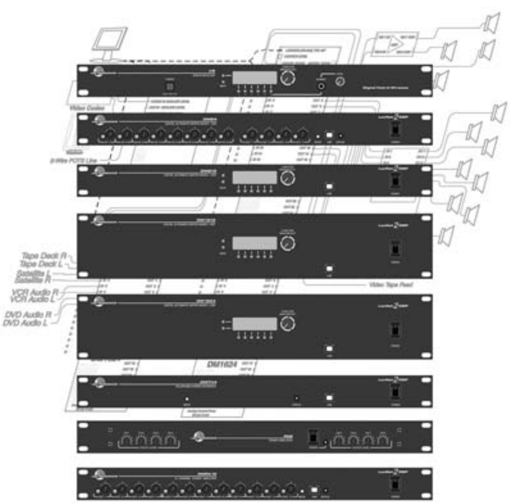


Figure 34-40. An automatic mixer with matrix output and DSP. Courtesy Lectrosonics.

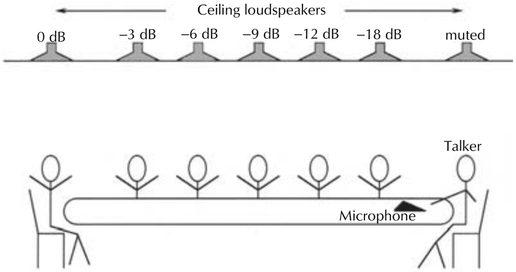


Figure 34-41. A simplified mix-minus system.

A mix-minus system is complex but the calculations are simple if taken one microphone at a time. Consider a single talker and microphone. Assume that the listener seated next to this talker can hear the talker unaided by the sound system. Now, use inverse-square law, Eq. 34-1, to calculate the loss from this nearby listener to the farthest listener. Amplify the signal to this farthest listener (using the loudspeaker nearest to that listener) to make up for the loss. Do the same for each remaining listener. Now, repeat the process for the second talker and so on. A spreadsheet is a useful tool for keeping track of these calculations and the required settings in the matrix.

34.4.3.4 Problems in Automatic Mixing

Despite sophisticated circuitry, ambient noise may still turn a microphone channel on at the wrong time. Another problem is coherent input signals, that is, signals that are in-phase and have similar waveshape, which may fool the mixer and allow it to raise its gain to a feedback condition. Nearly coherent signals may arrive at the microphones from a slammed door, for example.

An obvious problem with all automatic mixers is that they do not know when a new talker approaches the microphone. Thus, the mixer cannot readjust a microphone level for a loud-versus-quiet-voiced talker. A compressor or AGC (automatic gain control) circuit could be added to the mixer to adjust the level to compensate but this would defeat the number of open microphones (NOM) function and could cause the system to go into feedback. An experienced human operator is the best solution to this problem.

Despite their problems, automatic mixers are extremely useful, and a well-designed system with an automatic mixer is more likely than ever before to be audibly transparent to an audience.

34.4.4 Signal-Processing Components

Most signal-processing functions are now performed by software modules in a DSP device. However, it is still possible to purchase individual signal-processing devices and the functions are the same whether they are performed by a separate hardware device or a software module in a DSP. For reasons of clarity, this section considers separate hardware signal-processing components. However, as discussed, the functions can be performed equally well by multifunction DSP devices.

34.4.4.1 Compressors and Limiters

Compressors and limiters are devices that control a system's dynamic range (dynamic range is the difference in dB, between the highest and lowest L_P levels in any audio program). A limiter reduces the signal level when the level rises above a preset threshold. In this manner, a limiter helps minimize system damage from dropped microphones or other transients, Fig. 34-42.

A true compressor reduces too-high signal levels but it also increases very low signal levels to keep them above the ambient noise. Compression ratio is the ratio of output level change to input level change in dB.

Although true compressors are used in broadcast and recording applications, sound reinforcement systems seldom use true compressors. Perhaps for this reson the terms *compressor* and *limiter* are often used interchangeably in sound reinforcement literature.

34.4.4.1.1 Sound System Applications for a Compressor/Limiter

In a paging system, a true compressor can keep the average level of the voices of different announcers more constant so that paging can reach noisy areas of a factory or airport more consistently. In addition, because of reduced dynamic range, peaks are lowered, reducing the chance of clipping distortion.

In a large sound reinforcement system, such as a concert tour sound system, a limiter can reduce the chance of peak clipping and can thus help avoid amplifier or speaker damage from large turnon/turnoff transients or from sudden, loud feedback.

34.4.4.1.2 Problems with Compressor/Limiters

While useful, compressor/limiters are not cure-all devices. The compressor makes its decision to begin compressing by continuously monitoring the program level. Unfortunately, the highest levels are usually low bass notes. Thus, the compressor/limiter may compress the high frequencies needlessly when it detects a bass note that is too loud. One solution to this problem is to use a compressor on each output of an electronic crossover on a biamplified or triamplified system so that the compressor acts only on the frequencies in each band. Another solution is to use a separate compressor on each mixer input that may receive excessive program levels. Perhaps the best solution, for quality-conscious systems, is to use the limiter just to limit peaks. Set up the limiter with a high compression ratio and a high threshold so that it begins limiting only on potentially dangerous peaks and then limits them hard. With this setup, the limiter should be inaudible at normal program levels.

34.4.4.2 Equalizers

An *equalizer* is a device that controls the frequency response of a system or an individual source. An equalizer could be considered to be a large number of tone controls, operating at different frequencies, all in one device. Chapter 23 provides details of equalizers and their design.

There are two types of equalizer commonly used in sound reinforcement, the graphic equalizer and the parametric equalizer, Figs. 34-43 and 34-44.

34.4.4.2.1 Graphic Equalizers

Graphic equalizers usually have a series of slider-type controls that boost or cut each frequency. When the controls are adjusted up or down they resemble a graph of the unit's frequency response, hence the name *graphic equalizer*. Graphic equalizers are commonly available in 1 octave types or in ⅓-octave types. An octave-band equalizer has controls that are spaced 1 octave apart. A ⅓-octave-band equalizer has controls that are spaced ⅓ of an octave apart. Some manufacturers offer ⅙-octave spacing for part of the frequency range.

Octave-band equalizers are useful for adjusting the frequency response of an individual source. For example, to mellow the sound of a nasal-voice singer, use an octave-band equalizer connected to the insert points on the mixing console's input channel.

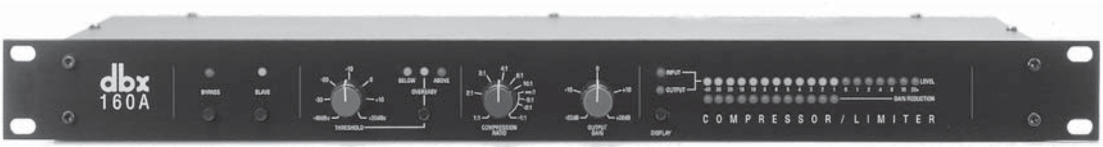


Figure 34-42. Compressor/limiter with threshold and compression ratio adjustments. Courtesy dbx Professional.

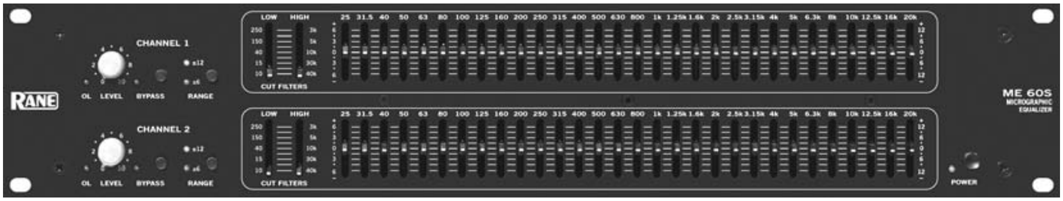

CHANNEL 1
CHANNEL 2
OL
LEVEL
BYPASS
RANGE
RANE
LOW
HIGH
CUT FILTERS
25 31.5 40 50 63 80 100 125 160 200 250 315 400 500 630 800 1k 1.25k 1.6k 2k 2.5k 3.15k 4k 5k 6.3k 8k 10k 12.5k 16k 20k
ME 60S
MICROGRAPHIC EQUALIZER
POWER

Figure 34-43. A two-channel analog graphic equalizer. Courtesy Rane.

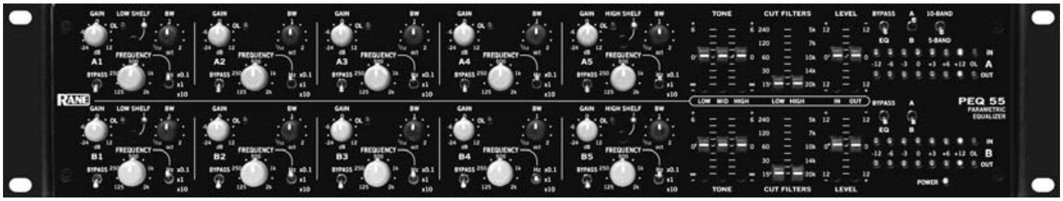


Figure 34-44. A two-channel analog parametric equalizer. Courtesy Rane.

For overall sound system equalization, as described in Section 34.5.2.2, an octave-band equalizer may be acceptable for simple systems in rooms with well-behaved acoustics. Most sound systems, however, need the greater precision of a ⅓-octave band equalizer or parametric equalizer as described below.

34.4.4.2.2 Boost-and-Cut Equalizers Versus Cut-Only Equalizers

Early equalizers were made from passive components and, thus, could not amplify a signal. These were cut-only equalizers. When active models were developed, however, they included electronic amplifiers for the purpose of buffering the impedance of the filters and, in some cases, of allowing the frequency response to be boosted as well as cut at any given frequency.

Either type of equalizer is suitable for sound reinforcement system equalization. However, choose a high-quality equalizer and when equalizing a sound system avoid boosting any frequency more than about +3 dB to maintain good system head room.

34.4.4.2.3 Constant Q versus Variable Q Equalizers

Q for a filter is the ratio between the filter's center frequency and its bandwidth. Early passive equalizers used variable Q filters. The Q of these filters was low at low insertion (small fader movement) and increased at high insertion. Some active equalizers have constant Q filters whose Q does not vary with insertion. Both types of filters can be combining filters. A good quality equalizer of either type is suitable for system equalization.

34.4.4.3 Parametric Equalizers

Parametric equalizers have fewer filters than graphic equalizers but the parameters of each filter are highly variable, hence the name, *parametric equalizer*. Typically, each filter of a parametric equalizer has variable insertion, variable Q and variable center frequency. Some mixing consoles include parametric equalization on each input. Others include quasi-parametric equalization where the insertion and center frequency, but not the Q, are variable.

Because of their flexibility, parametric equalizers with only three or four filter sections can approximate almost any curve needed for sound reinforcement equalization. For this reason, some designers favor them over graphic equalizers for sound system equalization.

34.4.4.4 Digital and DSP Equalizers

Most modern systems will have their equalization functions performed by some kind of multifunction DSP device where the equalization is simply a software module. Commonly, these DSP devices are controlled by a computer and the settings of an equalizer module are accessible via a user interface or GUI that resembles an analog equalizer of the same type. It may be possible to change many of the settings so the system designer can choose a graphic or parametric equalizer and control Q, center frequency, insertion depth and even filter design type. Usually the default settings are acceptable but it's a good idea to review all of these settings before using the device.

34.4.4.5 High-Pass and Low-Pass Filters

High-pass filters, also called low-cut filters, pass high frequencies and attenuate low frequencies. Low-pass filters, also called high-cut filters, pass low frequencies and attenuate high frequencies. A high-pass filter and low-pass filter with the same –3 dB frequency make a simple crossover network, Fig. 34-45.

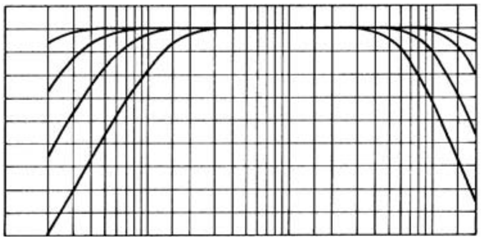


Figure 34-45. High-pass and low-pass filter curves.

Most sound reinforcement systems should include a 40 Hz, 18 dB/octave high-pass filter to reduce subsonic frequencies that might otherwise damage loudspeakers. For systems that include subwoofers, use a 20 Hz, 18 dB/octave (or steeper) high-pass filter in the subwoofer circuit and a 40 Hz high-pass filter (or higher) in the main loudspeaker circuit. Many graphic equalizers, and virtually all mult-function DSP devices, include a variable high-pass filter for this purpose.

34.4.4.6 Delay

Delay units are also called *signal delay* or *digital delay.* The term *time delay* is inappropriate since the signal is being delayed, not the time. Delay is useful in many sound reinforcement applications as detailed in Section 34.5.2.1. The most common example is a combination cluster and under-balcony loudspeaker system. The signal to the under-balcony loudspeakers is delayed to allow the sound from the cluster to catch up and avoid an artificial echo. Delay can also be used to line up the wavefronts from the high- and low-frequency components of a packaged loudspeaker system or to line up the wavefronts of the multiple loudspeakers in a cluster.

For sound reinforcement, choose a high-quality delay unit with dynamic range of 96 dB or greater and adjustment increments of 20 μs or shorter. For delaying the components of a packaged loudspeaker or a cluster, choose a delay with 10 μs or shorter increments.

Delay is a normal module of a multifunction DSP device. Most have varying increments of delay and a total delay limited only by system memory.

34.4.4.7 Electronic Crossovers

A crossover network routes high frequencies to the high-frequency loudspeakers (tweeters) and low frequencies to the low-frequency loudspeakers (woofers). The use of crossover networks is discussed in Section 34.3.4. Electronic crossovers should conform to the general specifications presented in Section 34.4.1 and are often included in multifunction DSP devices.

34.4.5 Digital Signal Processing

Most audio signal processing now takes place in the digital domain. Common DSP devices can be programmed to emulate a group of their analog counterparts arranged in a configuration chosen by the sound system designer. Some include mixing and output signal routing. There are two general types of digital signal-processing devices now in use in professional and commercial audio systems. (Also see Chapter 31.)

34.4.5.1 Multifunction DSP

The multi-function digital signal-processing system, typified by Peavey's Media Matrix, Biamp's Audia, or the BSS Soundweb, can become an entire sound system up to the power amplifiers and loudspeakers. This type of DSP device includes mixing and automatic mixing capabilities, output signal routing, and all types of signal processing (compression, limiting, equalization, delay, crossover), Fig. 34-46.

34.4.5.2 Power Amplifier DSP

Power amplifier DSP performs most of the same functions as all-in-one DSP. However the DSP devices are attached to the individual channels of a power amplifier. Because of this location, amplifier DSP is not suitable for mixing or output signal routing. However, it can perform all signal-processing functions and can also supervise and control the amplifier channel. Some power amplifier DSP devices are optional. Others are included with the amplifier and located inside the amplifier chassis, Figs. 34-47 and 34-48.

34.4.5.3 Loudspeaker Processing DSP

Most DSP functions needed by a loudspeaker can be performed by a multifunction or amplifier DSP. However, some manufacturers offer special-purpose DSP devices designed to perform loudspeaker optimization functions not available in these other devices. For

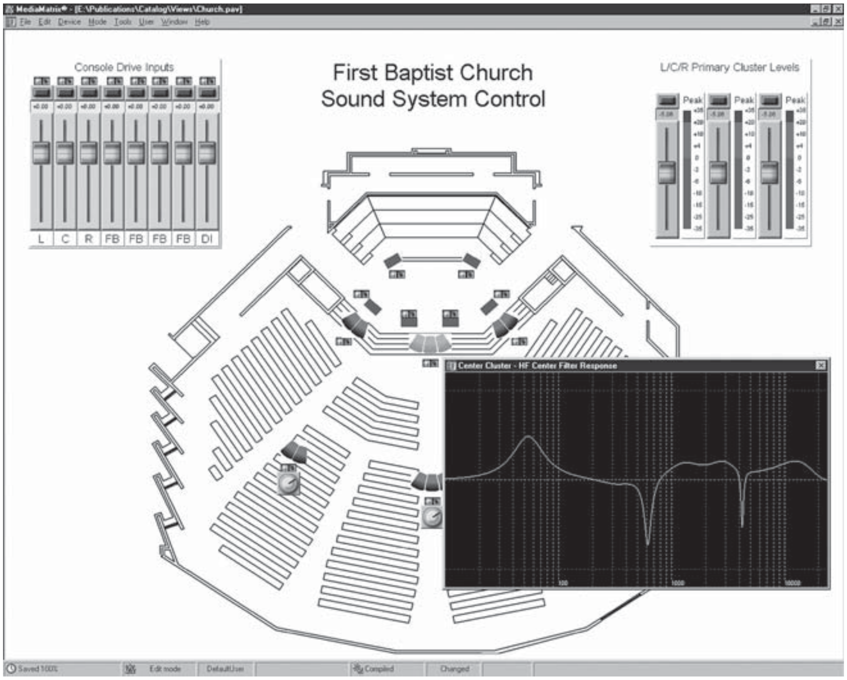

MediaMatrix® - [E:\Publications\Catalog\Views\Church.pav]
File Edit Device Mode Tools User Window Help
Console Drive Inputs
L C R FB FB FB FB DI
First Baptist Church
Sound System Control
L/C/R Primary Cluster Levels
Peak
Center Cluster - HF Center Filter Response
100
1000
10000
Saved 100%
Edit mode
DefaultUser
Compiled
Changed

Figure 34-46. Media Matrix. Courtesy Peavey.

QSC
DSP-4
POWER
SIGNAL
CHANNEL 1 INPUT
CHANNEL 2 INPUT
CHANNEL 1 OUTPUT
CHANNEL 2 OUTPUT
EXTERNAL POWER
15 V
300 mA
QSC AUDIO PRODUCTS INC. COSTA MESA CA

Figure 34-47. A two-channel power amplifier DSP. Courtesy QSC.

example, the loudspeaker processor could maintain a table of specific loudspeaker models with crossover, equalization, and delay settings for each model. Offloading this function from the multifunction DSP frees valuable outputs and software resources and may even reduce overall system cost, Fig. 34-49.

crown CDi1000
crown CDi2000

Figure 34-48. A power amplifier with built-in DSP. Courtesy Crown International.

34.4.5.4 Multifunction DSP in Sound Reinforcement System Design

DSP does more than simply replace traditional analog devices in a system design. DSP gives the system designer two important new capabilities.

First, DSP is programmable. That means the designer can set up more than one system design and give the user the ability to switch between “designs.” For example, the system could be equalized for both a full house and an empty house with a button push to switch between these curves. Or, the system at a religious facility could be optimized for a wedding or funeral or typical religious service with a button push to

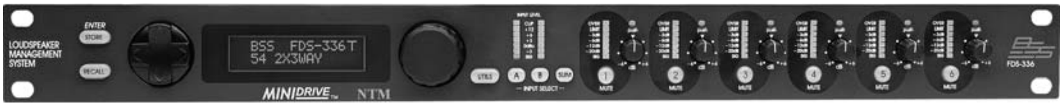


Figure 34-49. Loudspeaker DSP processor. Courtesy BSS.

select the appropriate configuration. This configuration can include equalization but it can also include mixing and output signal routing.

Second, DSP allows multiple signal-processing devices at reduced cost. This opens up opportunities that were previously available only in high-cost systems. For example, consider a religious facility with distributed column loudspeakers mounted on pillars down each side of the main auditorium. DSP can provide separate delay for each pair of loudspeakers, and it can even provide separate equalization. This can be valuable when the first pair is near the platform, the last pair is near a balcony, and the others are in yet a different acoustical environment.

34.4.6 Digital Audio Networking

More and more audio processing is done in the digital domain. Thus, it seems natural that audio signals should be transferred between devices in digital form. In some cases, this is a simple matter of connecting the digital output of one device to the digital input of another device. Audio devices with AES/EBU inputs and outputs are set up for this kind of connection.

In some systems, it may be useful to route multiple channels of audio from one location to another. For example in a performing arts center, it's commonly necessary to transfer multiple channels of audio from the stage to the mixing location and from there to the system rack room. If most processing is done in the digital domain it makes sense to keep the signals in digital form during these routing functions.

Although proprietary systems exist, most digital audio networking systems are based on Ethernet computer standards. The best known of these is CobraNet, developed by Peak Audio, a division of Cirrus Logic. CobraNet is licensed to a number of other manufacturers who have incorporated it into their products.

CobraNet and other digital audio networking systems enable multichannel digital audio transmission over CAT5 or fiber optic lines. They may also enable channel routing and patching functions, a sort of digital patch bay. Also, these transmission systems often include channels for system control signals and system monitoring information.

See Chapter 39 for a detailed discussion of digital audio networking.

34.4.7 Power Amplifiers

Power amplifiers for sound reinforcement should conform to the general specifications presented in Section 34.4.1. In addition, they must be able to drive professional loudspeaker loads and long loudspeaker lines. For this reason, a typical home entertainment power amplifier, while it may be a very high-quality product, is not suitable for professional or commercial usage, Figs. 24-50.

Figure 34-50. A two-channel professional power amplifier. Courtesy Crown International.

Most professional power amplifiers are two-channel, or multichannel solid state, analog devices with rack ears and cooling fans. Some include 70 V output transformers for use with distributed systems. Some have optional DSP modules as described in Section 34.4.5. Some have switching power suppliers to reduce their size and weight.

Power amplifiers should include an output relay or other method of uncoupling the loudspeakers from the power amplifier during turn on and turn off to avoid turn-on/turn-off transients from mixers and signal-processing devices. Also, the output relay disconnects the loudpeakers in the event of amplifier failure.

Some manufacturers offer multichannel power amplifiers with several power amplifiers in one chassis, Fig. 34-51. These power amplifiers can often be combined to form higher-power amplifiers or 70 V outputs. For multichannel systems, this type of power amplifier can often reduce costs.

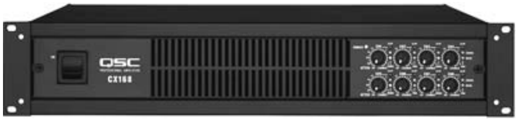


Figure 34-51. A multi-channel professional power amplifier. Courtesy QSC.

34.4.8 Pads and Transformers

A pad is a resistor circuit that reduces the output level from a source device to make it level compatible with a load device (also see Chapter 23). In the past, external pads were used to reduce the level from high-output microphones to make them compatible with normal microphone input circuits. Now, most mixers include "trim" controls to compensate for low- or high-level microphones.

Pads were also used to convert older high-level devices to be compatible with low-level devices. Today, most professional equipment has +4 dBu compatible input and output levels. In addition, passive components, like passive equalizers, are no longer in common usage. As a result, pads are usually only needed for the occasional connection of a professional device into a semipro or hi-fi device.

Transformers (also see Chapter 13) are devices that can be used to connect devices with unlike impedances and levels. For example, a hi-Z to lo-Z microphone transformer converts the high (voltage) level and high impedance of a high-impedance microphone to the low (voltage) level and low impedance of a low-impedance microphone input. Transformers can also be used to connect an unbalanced source to a balanced line. For example, a transformer could convert the unbalanced output of a consumer CD player to a balanced line for connection to a professional mixing console.

Loudspeaker transformers are used for 70 V loudspeakers as described in Section 34.4.8. Power amplifiers sometimes include transformers to convert the output of a conventional power amplifier to 70 V usage.

Transformers are level and impedance sensitive. That is, a microphone hi-Z to lo-Z transformer cannot be used for line-level impedance conversion. (It would distort.) Neither can a line-level transformer be used for microphone-level conversions. (It would also distort, although in a different manner.) Thus, when selecting transformers, needs must be defined in terms of both the impedance ratio desired and the level of the devices that will be connected to the transformer.

34.4.9 70.7 Volt /100 Volt System Design

A 70.7 V (referred to as 70 V) or 100 V loudspeaker system, as shown in Fig. 34-52, allows relatively long loudspeaker lines while minimizing I^2R line losses. 70 V or 100 V distribution also allows multiple loudspeakers to be connected to a single power amplifier without the need for complex series-parallel connections. For these reasons, 70 V or 100 V systems are commonly used for distributed ceiling loudspeaker systems and for any system where loudspeaker lines must be relatively long, 75 to 100 ft or more. Variations on this concept include 25 V distribution, which is sometimes used for school intercom systems, and 140 V distribution, which is sometimes used outside the United States. For brevity, the remainder of this discussion will use the term 70 V to refer to all of these systems. With the exception of local electrical codes, which may limit usage, design principles are the same for any of these systems.

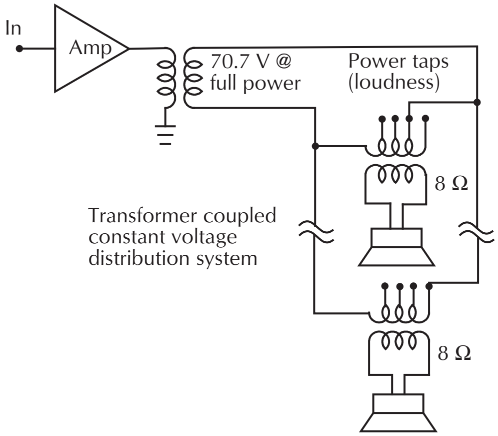


Figure 34-52. A 70.7 volt loudspeaker system. Courtesy Rane.

34.4.9.1 70 V Transformers

Fig. 34-53 shows a typical 70 V loudspeaker transformer. Choose a transformer that supports the impedance of the selected loudspeaker and has appropriate power taps for the application. Also choose a transformer with performance specifications that are appropriate for the application. In particular, pay attention to the transformer's frequency response, its distortion, and its dB loss figure. Low-cost 70 V transformers will have poor low-frequency performance, higher distortion (especially at low frequencies), and a dB loss of 1.5 dB or more. These specifications may be suitable for low-level paging and background music. Higher-cost 70 V transformers will have improved low-frequency response, reduced distortion at low frequencies and a dB loss of less than 1.5 dB. Use these for higher-performance applications.

Note that the dB loss specification is usually given in such a way that the transformer delivers its rated power to the loudspeaker but draws slightly more from the power amplifier. As an example, consider a high-quality

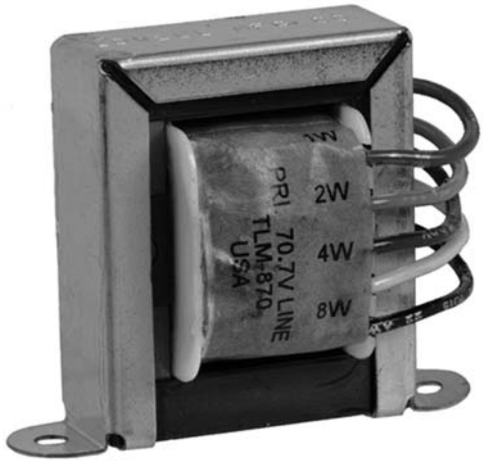


Figure 34-53. A typical 70 V transformer. Courtesy Lowell.

70 V transformer with a 0.5 dB loss. When this transformer delivers 10 W to the loudspeaker, it draws 11.1 W from the power amplifier. *This is very important in selecting the correct amplifier size.*

Choose the power level fed to the loudspeaker by connecting the appropriate primary winding to the 70 V line. Connect the loudspeaker to the appropriate impedance tap on the transformer secondary. Some manufacturers offer packages that include a ceiling loudspeaker with transformer preinstalled.

34.4.9.2 Designing a 70 V System

Choose the system loudspeakers and transformers. Calculate the power required per loudspeaker as described in Section 34.3.3.7.1. Then, choose an appropriate 70 V transformer. Next, calculate the power required by all of the system loudspeakers (the sum) and choose a power amplifier that's big enough to supply this power *plus the amount needed to overcome the loss in the transformers.* Here is an equation to help in this final calculation.

$$P_A = P_L 10^{\frac{L_T}{10}} \qquad (34\text{-}33)$$

where,

P_A is the required power amplifier size (minimum size in watts),

P_L is the total power required by all of the loudspeakers (the sum),

L_T is the loss of an individual 70 V transformer in dB (a positive number).

34.5 System Installation and Commissioning

34.5.1 What Is Commissioning?

After the system has been installed, it must be commissioned. Commissioning involves three general steps. These are outlined below and followed by a detailed discussion of selected commissioning topics.

34.5.1.1 Step 1—Test All Components

Test and confirm that all electronic components and all connections, including microphone and loudspeaker connectors and patch bays, are correctly wired and in proper working order. Pay special attention to polarity and other potential wiring errors.

34.5.1.2 Step 2—Adjust the Electronics

Set up system DSP to its final configuration. Do not adjust DSP delay, limiter or equalizer settings at this time. Next, set any loudspeaker DSP to its final configuration. If the loudspeaker DSP includes optimization for specific models of loudspeaker, implement this optimization at this time. Finally, adjust system gains and losses to minimize hum and noise and optimize head room.

34.5.1.3 Step 3—System Adjustments and Equalization

One at a time, adjust the system power amplifiers to produce the designed L_p in each audience area. Next, adjust any digital delays for satellite clusters or under-balcony loudspeakers. Finally, equalize the system as discussed in Section 34.5.2.2.

34.5.1.4 Connectors and Cabling

As simple a subject as this may seem, faulty connectors and cabling are the source of a majority of sound system problems. Well-made cabling, of the proper type, with the right connectors for the job, on the other hand, will keep a system operating at maximum efficiency with a minimum of noise pickup.

34.5.1.4.1 General Notes on Cable

A cable is a group of two or more wires, usually in a single outer (insulating) sheath, designed for a particular function.

Cables for portable audio systems should always be made from stranded, not solid, wire. Solid wire cables will break after the repeated flexing of portable usage. Shields should be braided wire, not foil, for the same reason. Cable for permanently installed systems, on the other hand, can utilize foil shields. In addition, while a tough, rubberized outer sheath is desirable for portable cable (like microphone cable), a smooth vinyl-type sheath will benefit the permanent system installer, since it pulls through conduit more easily.

34.5.1.4.2 General Notes on Connectors

There are only a few types of connectors in general use in commercial and professional sound systems, as shown in Figs. 34-54 and 34-55. The most common of these are discussed here.

34.5.1.4.3 XLR-Type Connection

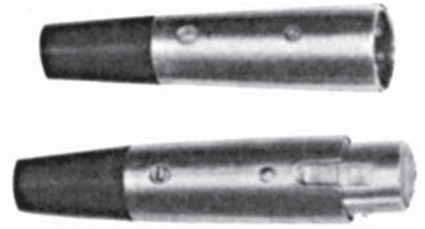
The term *XLR* was first used by the Cannon Company but has almost become a generic label for these high-quality audio connectors, now made not only by Cannon but also by Switchcraft, Neutrik, ADC, and others. XLRs are the connector of choice for microphones and any balanced low-level or line-level audio signal as well as AES/EBU digital connections.

34.5.1.4.4 Phone Plugs

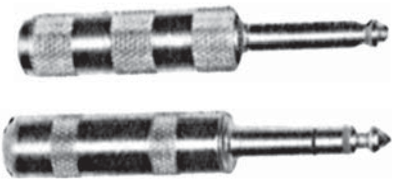
The term *phone* comes from the telephone industry, which normally used a type of phone plug in its early, nonautomated switchboards. Recording studio and other patch bays are close relatives of these telephone switchboards and often use a three-conductor variety of phone plug. The most common type of phone plug used in pro audio has a ¼ inch diameter shank and comes in two-wire (known as tip/sleeve, or T/S) and three-wire (known as tip/ring/sleeve, or T/R/S) versions. The ¼ inch phone plugs are commonly used for instrument amplifiers and hi-Z microphones and sometimes for portable loudspeaker connectors. Unlike XLRs, which are almost invariably high quality, the quality of commercially available phone plugs can vary widely.

34.5.1.4.5 RCA-Type Phono Plugs

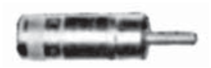
Note the term *phono* not phone, indicating that these plugs got their start on phonographs manufactured by the original RCA company. Phono plugs, or RCAs, are used primarily on hi-fi equipment but may be used to adapt a hi-fi tuner or cassette machine, for example, to an input of a professional mixer. Phono plugs, however, are fragile and do not make good general-purpose pro-audio connectors. Higher-quality phono plugs are used as the coaxial digital audio connectors on consumer equipment.



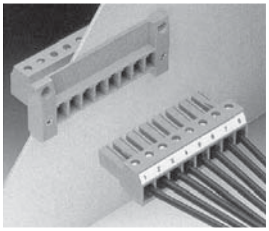
A. XLR.



B. Phone plug.



C. RCA (phono).



D. Miniature barrier block or Euroblock.

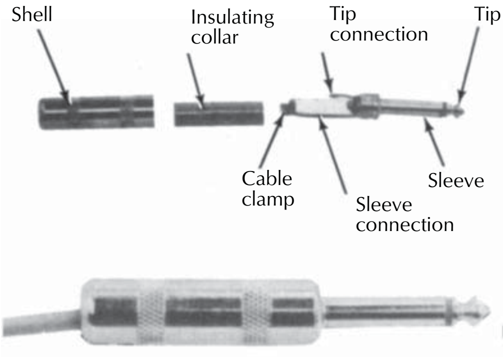
Figure 34-54. Various audio connectors.

34.5.1.4.6 Barrier Block Connectors

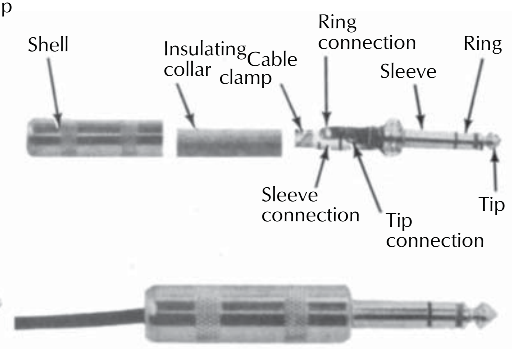
Professional audio products continue to get smaller while simultaneously adding more inputs and outputs. Manufacturers have responded by adopting miniature barrier block connectors, often called Phoenix connectors or Euro-Block connectors, for inputs and outputs. These connectors use a screw to capture individual bare wires in a small terminal hole.

34.5.1.4.7 CAT 5 Connectors

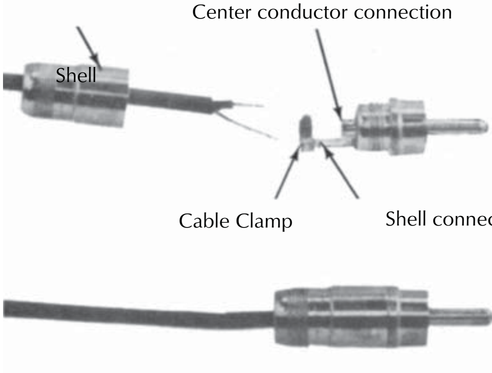
CAT5 connectors, also known as RJ45 connectors, are used for Ethernet networking and related digital audio connectors, Fig. 34-56.


Shell
Insulating collar
Tip connection
Tip
Cable clamp
Sleeve
Sleeve connection

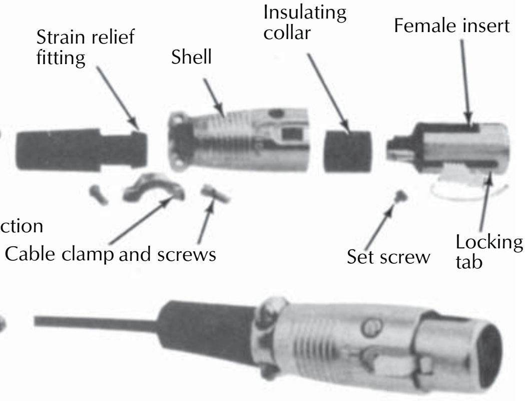
A. A standard two conductor phono plug.


Shell
Insulating collar
Cable clamp
Ring connection
Ring
Sleeve
Sleeve connection
Tip
Tip connection

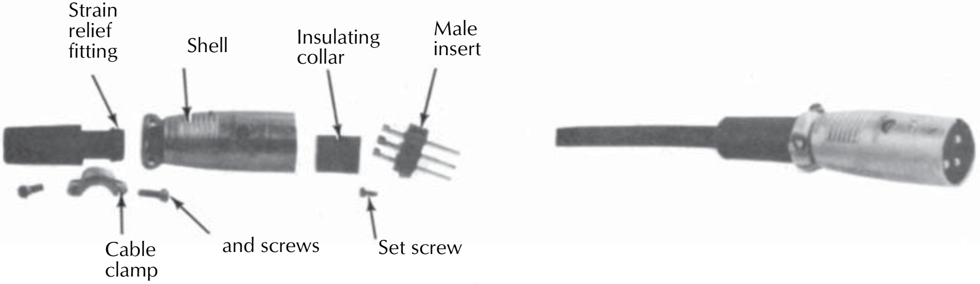
B. A tip, ring, sleeve, three conductor phone plug.


Center conductor connection
Shell
Cable Clamp
Shell connection

C. An RCA-type pin plug.


Strain relief fitting
Shell
Insulating collar
Female insert
Cable clamp and screws
Set screw
Locking tab

D. A female XLR connector.


Strain relief fitting
Shell
Insulating collar
Male insert
Cable clamp
and screws
Set screw

E. A male XLR connector.

Figure 34-55. Various audio connectors with parts identified. Courtesy Yamaha International Corp.

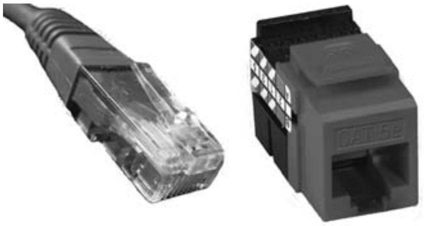


Figure 34-56. A CAT5 connector and socket.

34.5.1.4.8 Cable and Connectors for Microphones and Other Low-Level Devices

Lo-Z balanced microphones use shielded, two-wire cable and XLR-type connectors. Hi-Z (unbalanced) microphones usually use a ¼ inch phone plug connector. Exposed (portable) microphone cable should have a flexible, tough outer sheath; a braided shield; and stranded inner wires.

The XLR-type connector is an industry standard for lo-Z balanced microphones. Unfortunately, in the past, the wiring of these connectors was not completely standardized. Although pin 1 on the connector was almost always connected to the cable shield, some manufacturers used pin 2 as high or + and other manufacturers used pin 3 as high or + (with the remaining pin low or –). Today, most manufacturers use pin 2 as +. However, older microphones may still use pin 3 as +. Use of two microphones with different polarity to mic the same instrument or voice can result in undesirable phase cancellations. For this reason, it's wise to check the polarity of older products.

34.5.1.4.9 Microphone Snake Cables

A snake cable is actually a group of microphone or line-level cables all in one outer sheath. These cables use foil shields to reduce their overall diameter to a reasonable size. Because of the fragility of the foil shields in a snake cable (and because of the high cost per foot), extra care must be taken in their handling.

34.5.1.4.10 Cable and Connectors for Line-Level Devices

Line-level devices normally use the same type of cable and connectors as microphones and other low-level devices. That is, balanced line-level devices normally use XLR-type connectors and unbalanced line-level devices normally use ¼ inch phone plug connectors or RCA-type phono connectors. Some balanced line-level devices use three-conductor, ¼ inch tip/ring/sleeve (T/R/S) connectors.

Like older microphones, the polarity of XLR connectors on older line-level devices was, unfortunately, not standardized. Either pin 2 or pin 3 may be the + pin (pin 1 will almost always be the shield).

34.5.1.4.11 Cable and Connectors for Loudspeakers

Loudspeaker cable carries much higher levels of electrical power than either microphone or line-level cable. For this reason, loudspeaker cables use larger gauge wire. Typical loudspeaker cable uses anywhere from number 18 gauge wire to as large as number 10 wire (or even larger). Number 18 gauge wire is suitable only for low-level loudspeakers (like the hi-fi loudspeakers in your den). Number 16 gauge wire is suitable for short runs (less than 25 ft) of low- to medium-level pro-audio loudspeakers. Number 14 gauge wire is suitable for most pro-audio work unless loudspeaker runs are longer than about 75 ft. In that case, number 12 gauge wire should be used. For very long runs of high-power loudspeaker cable, use number 10 (or even number 8) wire. A better way to handle long loudspeaker cable runs, however, is to move the power amplifier closer to the loudspeakers and run line-level signals over the long distance. Alternately, use a 70 V (or 100 V) distribution system.

One apparent way to reduce loudspeaker cable requirements is to use powered loudspeakers that do not require *any* loudspeaker cable (only signal cable and ac power). However, unless ac power already exists at each loudspeaker location, the ac cable and conduit requirements may be more expensive than loudspeaker cable for nonpowered loudspeakers.

A connector developed by the Neutrik company, known as the Speakon, has become a de facto standard among most loudspeaker manufacturers. The Speakon connector, Fig. 34-57 is, in many ways, an ideal loudspeaker connector. It is a high-current twist-lock connector that is unlikely to fall out of its socket. It is self-wiping to keep its contacts clean. It is easy to assemble and is made from tough, lightweight plastic. In addition, it is relatively inexpensive in comparison to high-current metal connectors.

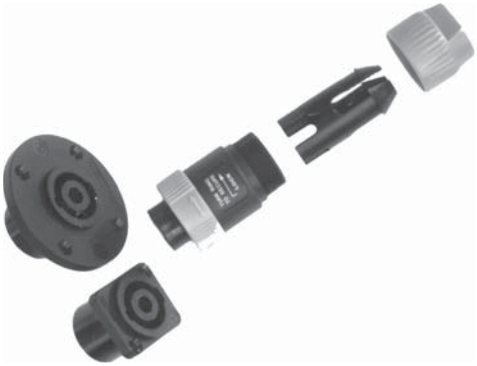


Figure 34-57. The Neutrik Speakon, a four or eight wire, high-current, twist lock loudspeaker connector. Courtesy Neutrik USA, Inc.

Except for very high-quality types, ¼ inch phone plugs are not suitable for the high-current use they get in pro audio. Thus, ¼ inch phone plugs are only suitable for low- and medium-level loudspeakers (perhaps up to 200 W or so per loudspeaker). Some power amplifier outputs use dual banana connectors, also called five-way binding posts. XLR connectors are sometimes used for loudspeaker connectors, but their current capacity, like the capacity of a phone plug, is limited, and they should not be used for higher-power capacity systems.

34.5.1.5 Cable and Connectors for Digital Audio Devices

Cable and connectors for digital audio devices are derived from computer cabling and are often exactly the same. For example, many digital audio devices utilize Ethernet-style CAT5 cabling and connectors or USB cable and connectors. As a rule, these computer-derived connectors and cabling are not rugged enough for portable sound system usage. For this reason, some audio cabling companies have introduced specialty connectors designed for portable usage such as the one shown in Fig. 34-58.

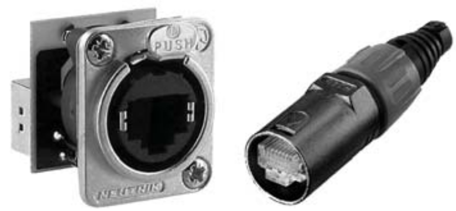


Figure 34-58. An ethernet-style connector designed for portable audio systems. Courtesy Neutrik USA, Inc.

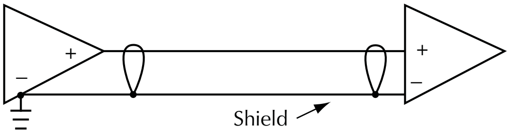
The AES/EBU digital audio standard utilizes conventional XLR connectors. The digital audio output coaxial connector found on home-theater receivers is a high-quality RCA-type connector. As with all computer connections, digital audio connections should use high-quality cable and connectors of the right impedance and length. Consult the device manufacturer for recommended cable specifications.

Fiber optic cable and connectors may be used in large installed audio systems because of their ability to carry multiple channels of digital information (audio, video, control, and other) on a single cable and because of their relative immunity from hum and noise pickup. Fiber optic suppliers often provide seminars to teach designers and installers how to specify, design, and install fiber optic cabling systems.

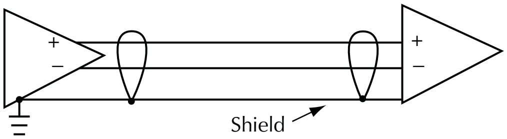
See Chapter 15 for a detailed discussion of fiber optic cabling systems. See Chapter 39 for a detailed discussion of digital audio networking and the associated connections.

34.5.1.6 Understanding Balanced and Unbalanced Lines

Every audio signal requires at least two wires. In an unbalanced line, the shield (outer conductor) is also one of the audio signal wires. Thus, an unbalanced line, Fig. 34-59A, needs only the shield and one additional wire (a total of two wires). In a balanced line, the shield does not carry audio signals. Thus, a balanced line, Fig. 34-59B, requires the shield and two additional wires to carry the audio signal (for a total of three wires). Also see Chapter 37 for a more detailed discussion of balanced and unbalanced circuits.



A. An unbalanced line connecting two audio devices.



B. A balanced line connecting two audio devices.

Figure 34-59. Balanced and unbalanced lines. Courtesy Fender Musical Instruments.

The primary advantage of a balanced line is that it is much less likely to pick up external electronic noises (hum, buzzing, static, radio stations, etc.) than an unbalanced line.

34.5.1.7 Impedance and Level Watching

While some passive devices require impedance matching, most active audio devices do not require matched impedances. What they do require is impedance compatibility. In addition, all audio devices require signal level compatibility. Thus, impedance and level watching means establishing and maintaining required impedance and level compatibility (as will be discussed).

34.5.1.7.1 Terms: Source, Input, Output, Load

In Fig. 34-60, the source is the microphone, the input is the input to the mixer/amplifier, the output is the output from the mixer/amplifier, and the load is the loudspeaker, but these four terms are relative. For example, the input to the mixer-amplifier can be called a load from the viewpoint of the microphone. And, the mixer-amplifier output can be called a source from the viewpoint of the loudspeaker.

Thus, the input impedance of the mixer/amplifier can be called the load impedance for the microphone, and

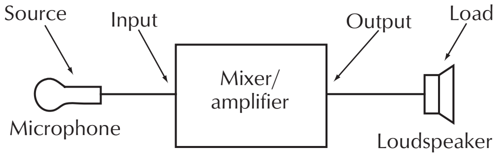
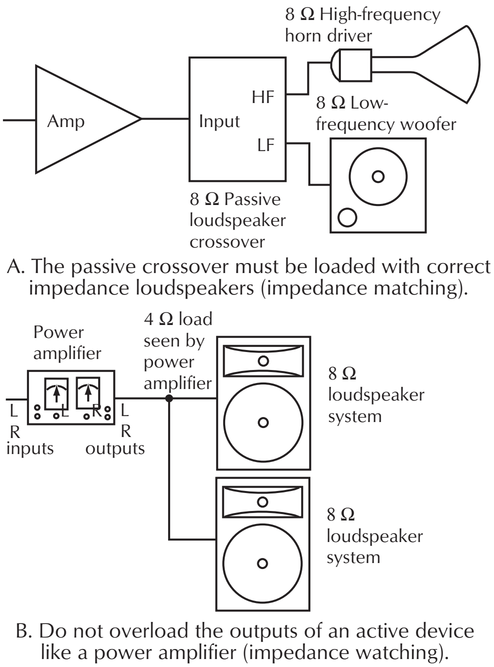


Figure 34-60. Origin of terms. Courtesy Fender Musical Instruments.

the output impedance of the mixer/amplifier can be called the source impedance for the loudspeaker.

These four terms—source, input, output, and load—and their relative nature are important to an understanding of impedance and level watching. As an example, consider a microphone whose impedance is 200 Ω. That impedance is actually the internal impedance of the microphone and should be called the source or output impedance of the microphone. (The microphone is a source from the viewpoint of the mixer/amplifier.)

That same microphone should probably be loaded with an impedance of 2 kΩ or higher. That load impedance is actually the input impedance of the mixer/amplifier (the input of the mixer/amplifier is a load to the microphone).



A. The passive crossover must be loaded with correct impedance loudspeakers (impedance matching).

B. Do not overload the outputs of an active device like a power amplifier (impedance watching).

Figure 34-61. Impedance watching. Courtesy Fender Musical Instruments.

34.5.1.7.2 Impedance Compatibility

Impedance watching just means making sure that when two devices are connected, they are compatible from an impedance viewpoint, Fig. 34-61. Here are some rules to help watch impedances.

34.5.1.7.3 Passive Devices

In the special case of a passive filter, like a loudspeaker crossover network or a passive graphic equalizer, input and output impedances must be matched. These devices are the origin of the familiar term *impedance matching*. Impedance matching means that if the device is a loudspeaker crossover network and it has an 8 Ω low-frequency output impedance and an 8 Ω high-frequency output impedance, then it must be connected to an 8 Ω low-frequency loudspeaker and an 8 Ω high-frequency loudspeaker. Any other impedance, either higher or lower, will degrade the performance of the crossover network. (The input to a modern loudspeaker crossover network is designed for the very low actual output impedance of a modern power amplifier.)

An older passive device such as a passive graphic equalizer has similar requirements. If such a device has a 600 Ω input impedance, then it must be connected to a source impedance of exactly 600 Ω. The same goes for the output. If the passive graphic has a 600 Ω output impedance, then it must be connected to a load impedance of exactly 600 Ω to insure proper operation of the graphic equalizer. In many cases, build-out and termination resistors must be added to match these impedances (see Chapter 23).

34.5.1.7.4 Passive Sources

Impedance matching for a passive source, like a dynamic microphone or guitar pickup, simply means supplying a compatible load impedance for that device. The device specifications should be a reasonably accurate guide to the proper load impedance. A good rule of thumb for dynamic microphones is that the microphone load impedance (which is probably the input impedance of a mixer or preamplifier) should be at least five to ten times the microphone's rated source impedance. Thus, for a 150 Ω (source impedance) microphone, the optimum load impedance would be 750–1500 Ω or higher. This requirement is satisfied by the input of almost all low-impedance mixer inputs. Note that the load impedance required by a high-impedance microphone is many times higher than the load impedance required by a low-impedance microphone. High-impedance microphones, therefore, can only be used with mixers having special inputs designed for these high impedances.

34.5.1.7.5 Active Devices

An active device is one that uses batteries or ac power and has one or more tubes, transistors, or ICs. Impedance watching for an active device means not overloading its output, that is, not connecting too low a load impedance to the output of the active device. A too-low impedance is an overload because the lower the impedance, the closer it is to a short circuit.

It's usually very easy to follow this rule because almost every active device comes with a set of specifications that indicates the value in ohms of the lowest allowable load impedance. This is usually called the rated or minimum load impedance. Incidentally, in almost every case, it's acceptable to connect a higher than rated load impedance to any active device.

For many modern solid state power amplifiers, for example, the minimum load impedance is 4 Ω. That means any impedance down to 4 Ω may be connected to this power amplifier. Since an 8 Ω loudspeaker is greater than 4 Ω, it is an acceptable load; a 16 Ω loudspeaker is also acceptable. Two 8 Ω loudspeakers in parallel equals a 4 Ω load so this arrangement is also acceptable. Four 4 Ω loudspeakers in parallel equal a 1 Ω load; this is definitely not acceptable. Connecting a too-low load impedance to a power amplifier will cause the protection circuits of the power amplifier to operate, which increases distortion, and may, in extreme cases, cause damage to the power amplifier or loudspeakers.

For a line-level active device, like a limiter, the same rule applies. If the limiter has a rated minimum load impedance of 600 Ω, the output of the limiter may be connected to the input of any device whose input impedance is 600 Ω or higher. (The input impedance of most active devices is considerably higher than 600 Ω)

Some professional power amplifiers, on the other hand, have input impedances of 5 kΩ or lower. Connecting a hi-fi-type tuner, with a 10 kΩ minimum load impedance to the professional power amplifier, with its 5 kΩ input impedance would reduce the output level from the tuner and might also cause an increase in distortion.

34.5.1.7.6 Active Sources

Active sources like battery or phantom-powered condenser microphones should receive the same treatment as any other active device although most battery or phantom-powered microphones are designed to act like conventional low-Z dynamic microphones from the point of view of their desired load impedance.

34.5.1.7.7 Impedance and Cable Length

One more aspect of impedance watching involves the effect of cable length on the frequency response of high-impedance microphones. From the following information, we can see that a too-long cable on a high-impedance microphone will cause a loss in high-frequency response; that is, the sound from the microphone will be dull, and voices will lack intelligibility. This results from the interaction between the capacitance in the cable and the high impedance of the microphone, which form a low-pass filter. The lower impedance of a low-impedance microphone also interacts with the capacitance of the cable, but the effect is noticeable only at very high frequencies (out of the audio range). A good rule of thumb is to avoid cables longer than 15 ft with a high-impedance microphone (some high-impedance microphones will tolerate cable lengths up to about 25 ft). A low-impedance microphone, on the other hand, will perform properly with cables as long as 200 ft or more.

This same cable length consideration applies to line-level devices. The hi-fi tuner mentioned previously, for example, should not be used with a cable longer than about 15 ft. (The cable should be shorter if possible.)

34.5.1.7.8 Signal-Level Compatibility

Achieving level compatibility between devices means two things: avoiding too-high levels, which cause clipping distortion, and avoiding too-low levels, which allow electronic noises (usually hiss), as shown in Figs. 34-62A and 34-62B.

There are three basic classifications of level in analog professional audio devices, Fig.34-63:

1. Low-level devices (microphones, pickups, and so on).
2. Line-level devices (limiters, graphic equalizers, and so on).
3. High-level devices (the output from a power amplifier).

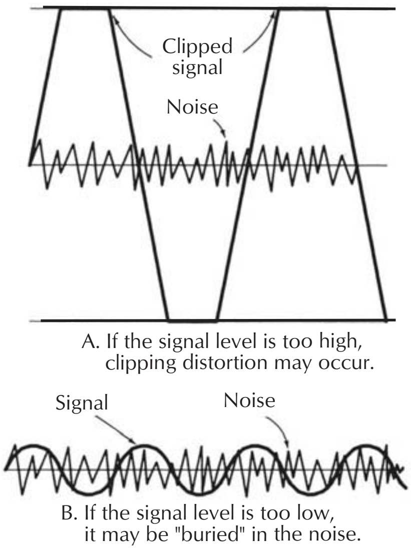

Clipped signal
Noise
A. If the signal level is too high, clipping distortion may occur.
Signal
Noise
B. If the signal level is too low, it may be "buried" in the noise.

Figure 34-62. Level watching. Courtesy Fender Musical Instruments.

The first rule of level compatibility is to avoid connecting devices from different classifications unless they are specifically designed for each other.

For example, don't connect a microphone directly to a power amplifier because the output of the microphone is too low. This connection wouldn't damage anything but would result in very low sound level, and the noise from the power amplifier might be almost as high as the wanted sound.

As an obvious example, don't connect the output from a power amplifier to the input of a mixer. The power amplifier output level is far too high for the input of the mixer. This connection would almost certainly result in severe clipping distortion (the mixer might even be damaged).

Many devices, however, have an input that is compatible with one level and an output that is compatible with the next higher level. For example, the input channels of most professional mixers are compatible with low-level devices like microphones, although their outputs are designed for both mic-level and low- and high-level line loads. A power amplifier, as another example, has a line-level input and a loudspeaker-level output. Thus, the output of a limiter or graphic equalizer can usually be connected directly to the input of a professional power amplifier. (Impedance must be considered, too, but the input impedance of most professional power amplifiers is high enough to be impedance compatible with the output of almost any professional line-level device.)

The situation is complicated somewhat by variations in the level of devices in a given category. For example, a typical condenser microphone has a higher output than a typical dynamic microphone. One solution to this problem is to design the mixer for the lower-level microphone and provide a pad for the higher-level microphone. A more common solution on professional mixers is to include either a built-in (adjustable) pad or a preamplifier gain adjustment or both. By properly adjusting these controls, the mixer's input channel can be optimized for either a dynamic or condenser microphone (and, with some mixers, for a line-level input).

The same kind of level-compatibility problems show up in line-level devices. Some line-level devices, mostly special effects devices, are designed for input and output voltages as low as −20 dBu. Others, including some tape machines, are designed for input and output voltages of −10 dBu. Most professional line-level equipment, however, is designed for input and output voltages of +4 dBu.

The level in dBu is found from

$$dBu = 20\log\frac{V}{0.7745} \tag{34-34}$$

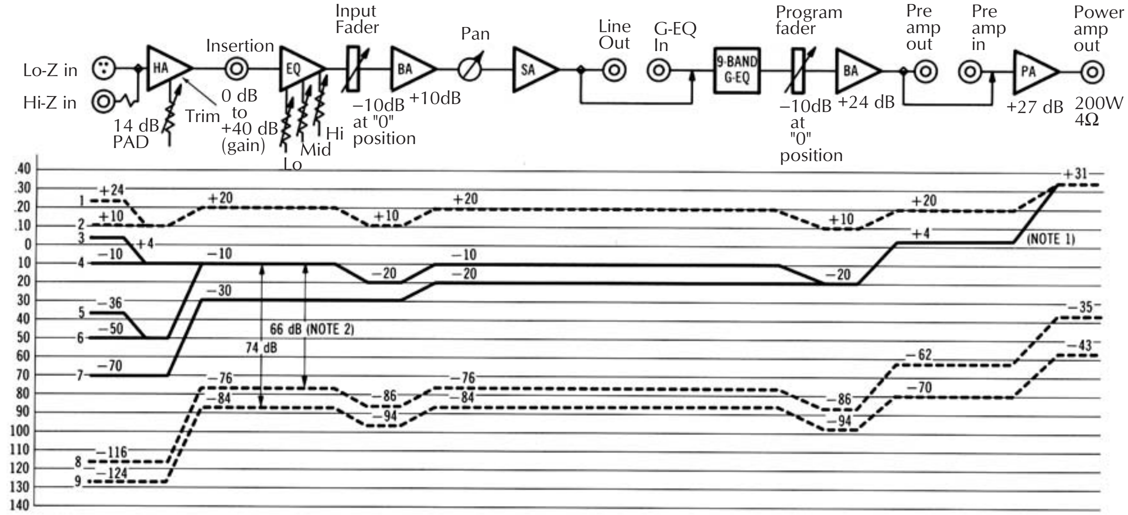
where,

dBu is the voltage level in dBu,
V is the voltage level in volts,
0.7745 is the reference level for dBu in volts.

dBu specifications are only used for voltage output ratings and are not the same as dBm ratings. Many manufacturers are now rating the input and output levels of their products in dBu, however, so it is useful to understand this specification.

The process of achieving compatibility with these line-level devices is similar to the process for low-level devices (e.g., the microphones discussed earlier). Whenever possible, connect the output of a −20 dBu device to the input of another −20 dBu device (the same applies to −10 dBu devices and +4 dBu devices).

If this isn't possible and the source device has a higher output level than the load device, use a pad to attenuate the level of the source device. For example, if the source is a +4 dBu limiter, and the load is the input to a −10 dBu tape machine, a 14 dB pad is needed to achieve level compatibility. To design a proper pad, impedances must be taken into account (see Chapter 23). Without the pad, there is a risk of clipping distortion. Just turning down the output of the source device probably won't solve the problem, either. This may result in that other level-compatibility problem—electronic hiss noise.



1 Hi-Z in clip line
2 Lo-Z in clip line
3 Hi-Z in nominal line (trim at min)
4 Lo-Z in nominal line (trim at min)
5 Hi-Z in nominal line (trim at max)
6 Lo-Z in nominal line (trim at max)
7 Lo-Z in max gain line (trim & all faders at max)
8 Hi-Z in noise floor line
9 Lo-Z in noise floor line

Note 1: +4 dB at the power amp in jack produces 200 W/4 Ω at the power amp out jack.
Note 2: Signal-to-noise ratio of lo-Z in is 74 dB; of hi-Z in is 66 dB.
Note 3: 0-dB reference is 0.775 V.

Figure 34-63. Typical sound reinforcement system gain/level chart. Courtesy Fender Musical Instruments.

If the source device has a lower output level than the load device, a line-level preamplifier could be connected between them to give the required amount of gain. In many cases, however, simply connecting these two devices will prove satisfactory. The worst that can happen here is additional hiss, which may be tolerable.

34.5.1.7.9 The dBV Specification

All professional equipment uses a dBu or dBm reference for any "+4" input or output. Some professional microphones and consumer type equipment, however, may have outputs that are rated in dB using a dBV reference. When in doubt about the dB reference, consult the owner's manual for the equipment or, measure the output with a known signal.

The level in dBV is found from

$$dBV = 20\log\frac{V}{1} \tag{34-35}$$

where,

dBV is the voltage level in dBV,

V is the voltage level in volts,

1 is the reference level for dBV in volts.

34.5.1.7.10 Impedance and Power Transfer

To understand what happens to the power output of an amplifier when different impedances are connected to it, find the rated power output of the amplifier and its rated load impedance. That rated load impedance is usually the minimum acceptable load impedance of the amplifier.

In addition, find the true minimum impedance of the loudspeaker as well as its rated or nominal impedance. Normally, the nominal impedance of the loudspeaker will be used to make impedance-matching calculations like those described in the next paragraph. The minimum impedance of a loudspeaker, however, can fall significantly below its nominal impedance, and a loudspeaker with an extremely low minimum impedance could even overload a power amplifier whose rated load impedance was acceptable for rated nominal impedance of the loudspeaker.

Some professional power amplifiers are designed to accept impedances as low as 2 Ω because of the very low minimum impedance of some loudspeakers. Many 8 Ω loudspeakers (8 Ω is the rated or nominal impedance), for example, have minimum impedances of 6 Ω or even as low as 5 Ω. Two of these loudspeakers in parallel would have a minimum impedance of 2.5 Ω, which would still be within the safe limits for these power amplifiers.

The easiest way to describe the change in power output with different load impedances is to take an example, such as is shown in Fig. 34-64. One manufacturer's professional power amplifier is rated at 440 W per channel into a 4 Ω load. It has a minimum load impedance of 2 Ω even though its 440 W power rating is at 4 Ω, and 440 W into 4 Ω means exactly that. Connect a 4 Ω loudspeaker to one channel of this power amplifier, and the amplifier will produce as much as 440 W into that loudspeaker. Connecting two 8 Ω loudspeakers (in parallel) to one channel of this power amplifier will, again, result in as much as 440 W into the resulting total impedance of 4 Ω. Each 8 Ω loudspeaker in this example will receive exactly one-half of the total power, or a maximum of 220 W.

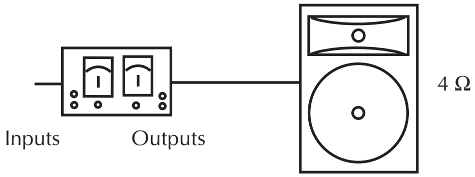
Connect a single 8 Ω loudspeaker to one channel of this power amplifier, and that loudspeaker will still only receive a maximum of about 220 W. (The actual power will be slightly higher.) Connect a single 16 Ω loudspeaker, and it will receive a maximum of about 110 W. In other words, doubling the load impedance halves the power output of a power amplifier. Conversely, halving the load impedance doubles the power output of the amplifier. Remembering this simple relationship can help insure that a loudspeaker and power amplifier will be compatible in terms of impedance and power levels.

34.5.1.8 Digital Audio Level and Impedance Watching

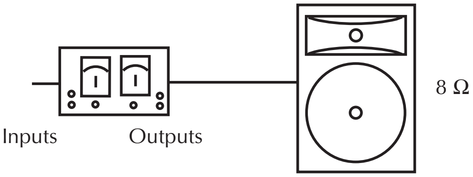
Since most digital audio connections and cabling are based on computer standards, an audio system designer or installer may confidently use properly rated computer cable and connectors to connect digital audio devices. Pay attention to maximum cable length and, for portable applications, use connectors and cabling designed for portable usage. For proprietary or noncomputer-related digital audio connectors, such as the AES/EBU connection, consult the device manufacturer for cable and connector recommendations. Also see Chapter 39 for a detailed discussion of digital audio networking and connections.

34.5.1.9 Grounding and Shielding (also see Chapter 32)

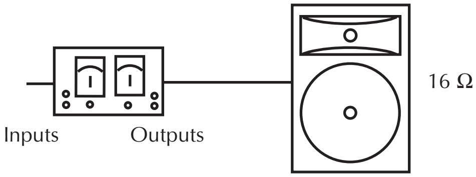
Caution: In any audio system installation, governmental and insurance underwriters' electrical codes must be observed. These codes are based on safety and may vary in different localities. In all cases, local codes take precedence over any suggestions contained in the *Handbook for Sound Engineers*.



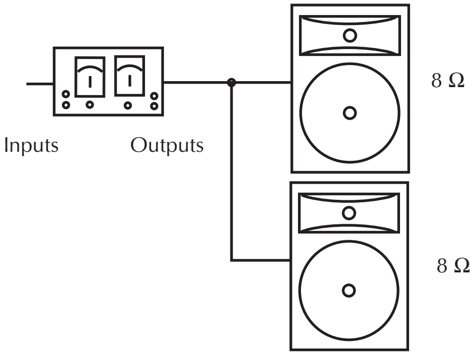
A. One 4 Ω loudspeaker draws a maximum of 440 W from the amplifier.



B. One 8 Ω loudspeaker draws a maximum of 220 W from the amplifier.



C. One 16 Ω loudspeaker draws a maximum of 110 W from the amplifier.



D. Two 8 Ω loudspeakers connected in parallel draw a maximum of 440 W total or 220 W each from the amplifier.

Figure 34-64. Impedance and power transfer. Courtesy Fender Musical Instruments.

Note that the ac power discussions in this section apply specifically to the United States only. The general discussions of grounding and shielding, however, should be applicable to audio systems used in any location. Always obey local and national fire and electrical safety regulations.

There are two primary reasons for careful grounding and shielding in an audio system. The first reason is safety. A poorly grounded system, especially outdoors, may be a shock hazard. The second reason is to reduce pickup of external noise. That external noise expresses itself in the form of hums and buzzes and other noises including radio station pickup.

34.5.1.9.1 Grounding for Safety

In U.S. electrical systems, the third (round) prong on the ac cable of any piece of audio equipment is the ac safety ground. When plugged into a properly wired ac receptacle, the third prong of the ac cable connects the chassis of the audio device to the ac ground through the third prong of the ac receptacle.

This is the ideal situation from a safety viewpoint. Under these conditions, there is almost no combination of events that could cause a shock hazard from a single audio device by itself. It is unfortunate (from a safety viewpoint) that any audio device is seldom used by itself; there are always other pieces of equipment involved, and most of the time, these are also ac powered. In addition, one or both of the following may be encountered:

1. Older audio equipment (in particular, guitar amplifiers) with two-wire ac plugs and ground or hum switches.
2. Older, two-wire ac receptacles or improperly wired three-wire ac receptacles.

34.5.1.9.2 Improperly Wired ac Receptacles

An improperly wired outlet, Fig. 34-65, may have its two ac wires reversed (polarity reversal), or it may have a disconnected ground. Any fault in the wiring of the ac receptacle is potentially hazardous, and, thus, the best, and perhaps the only safe way to deal with an improperly wired ac receptacle is to simply refuse to use it until it has been properly repaired by a licensed electrician.

A simple three-prong outlet tester can indicate many of these problems and is a useful addition to any audio technician's tool kit. *Note that the three-prong outlet tester cannot detect a ground-neutral swap. Neither can it detect a high-impedance ground. Both of these are hazardous conditions.*

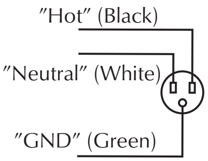
Another worthwhile measurement is the actual ac receptacle voltage, especially in an unfamiliar facility. Voltages that are too high or too low could cause improper operation, or even damage the equipment; *too-high voltages could also pose a shock hazard.* Most audio equipment will work fine on an ac outlet with voltages as low as about 110 Vac and as high as about 120 Vac. Newer equipment may be designed to automatically adjust for voltages as low as 100 Vac and as high as 240 Vac. Check the specifications for the equipment if there is any doubt.

34.5.1.9.3 Two-Wire ac Receptacles

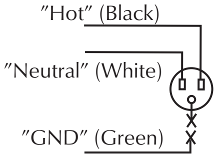
All new ac installations in the United States use modern three-wire ac receptacles with a third ground prong. The problem with two-wire ac receptacles is that they don't have that important third ground prong. Thus, to use one of these two-wire receptacles, it's necessary to adapt it to the three-wire ac plug on a more modern piece of audio equipment using a two-wire to three-wire ac adapter. When the two-wire ac outlet is wired properly and a low-impedance grounded screw is available, these adapters can maintain a safe ground for the three-wire audio equipment.

To make this two-wire adapter work properly, connect the loose (ground) wire with a connector or the ground lug of the adapter to a grounded screw on the two-wire ac receptacle. To check the safety of the two-wire to three-wire connection, first, connect the loose wire on the adapter to the screw on the two-wire receptacle; then plug the two-wire adapter into the two-wire receptacle. Next, plug a three-wire ac outlet tester into the adapter. If the screw is grounded, the ac outlet tester will so indicate. (Most three-wire ac outlet testers either have a "good" light or they don't light at all on a good receptacle.) If the screw is not grounded, the outlet tester will so indicate. In this case, connect the loose wire from the adapter to some other grounded screw in order to maintain a safe ground. If the outlet tester shows a good ground but reversed polarity on the two-wire to three-wire adapter, simply reverse the adapter in the receptacle.

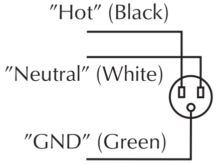
Remember that the three-prong outlet tester cannot detect a ground-neutral swap, a hazardous condition. Also, it is possible for the three-prong outlet tester to indicate a good ground when the ground connection is actually a high-impedance ground (poor connection). *For these reasons, it is strongly recommended that, whenever you are connecting any three-wire ac-powered equipment to any two-wire ac receptacle, you use a portable GFI (ground fault interruption) device to protect the equipment and the users.*



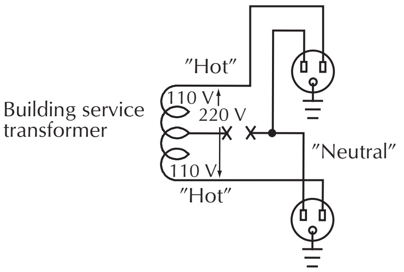
A. Properly wired 110 Vac outlet.



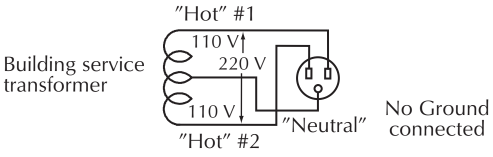
B. 110 Vac outlet with disconnected – ac ground wire creating potential shock hazard.



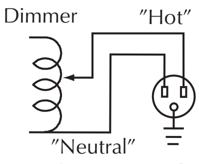
C. 110 Vac outlet with polarity (hot and neutral) reversed creating shock hazard and causing possible noise.



D. 110 Vac outlets with open neutral. Outlets will operate with voltage varying from 0 to 220 Vac, creating shock hazard and causing possible equipment damage.



E. 110 Vac outlet with a 220 Vac circuit connected to it. This is a highly dangerous and illegal connection.



F. A 110 Vac outlet connected to a light dimmer circuit, a dangerous and illegal connection.

Figure 34-65. Ac receptacle problems. Courtesy Yamaha International Corp.

34.5.1.9.4 Older Two-Wire Audio Equipment

Some newer equipment, especially consumer-type equipment, may come with a two-wire ac cable. This newer equipment may be as safe as if it had a three-wire (grounded) ac cable. A good example of such a piece of (nonaudio) equipment with a two-wire ac cable is a double-insulated power tool (drills, saws, and so on). One way to judge the safety of a piece of audio equipment with a two-wire ac cable is to look for a UL (Underwriter's Laboratories) sticker. Listings from other recognized safety agencies may also be used to judge the safety of a piece of equipment.

It's the older, two-wire audio equipment, however, that can be potentially hazardous. The details of how a shock hazard can develop are complex, but dealing with this problem in an audio system is straightforward. The shock hazard, if there is one, will probably develop between the chassis of an older, two-wire guitar amplifier and the chassis of a microphone.

The chassis of the microphone is connected to the chassis of the microphone-preamplifier-mixer through the shield of the connecting cable. Thus, if the mixer is properly grounded, the chassis of the microphone is properly grounded, too, and neither the microphone nor the mixer will present any safety hazard. The guitar amplifier (or other two-wire equipment), however, is, potentially, not properly grounded. That means that a hazardous ac voltage could be present on the chassis of the guitar amp or on the strings of a guitar, which are connected to the chassis of the amplifier through the shield of the guitar cord.

Although it is possible to test for this problem, it's also important to protect the user with a portable GFI (ground fault interruption) device on the guitar amplifier.

34.5.1.9.5 Grounding for Safety Outdoors

The most common safety problems outdoors are improperly wired portable ac outlets and wet ground or wet portable stages (and, of course, rain). Check the wiring carefully, using the same techniques as if the system were indoors. Consider canceling a performance if rain begins. If a performance must proceed on wet ground or in the rain, the best way to avoid shock hazards to the performers is to use wireless microphones and wireless instrument transmitters. These same outdoor problems, of course, can develop indoors on a damp floor. *Portable GFI (ground fault interruption) devices are strongly recommended for any equipment used in an outdoor or damp indoor situation.*

34.5.1.9.6 *Grounding to Reduce External Noise Pickup*

After safety, the second reason to pay attention to grounding is that, although proper grounding won't always reduce external noise pickup, poor grounding can unquestionably increase external noise pickup.

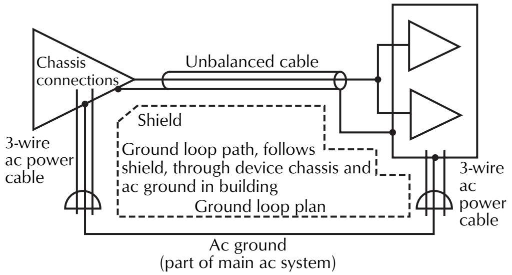
One myth about grounding is that a piece of equipment must be earth grounded to avoid noise pickup. Anyone who owns a portable MP3 player or CD player knows that this isn't true. Good grounding practice is primarily a matter of proper connections between devices, avoiding ground loops and using equipment that does not have the "Pin 1 Problem."

34.5.1.9.7 *Definitions: Ground, Earth, Common*

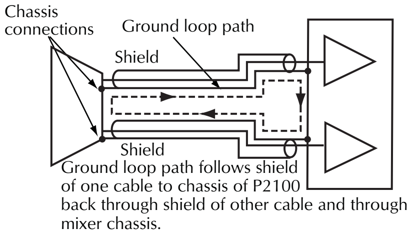
A *common connection* in an audio system is some point where a group of circuits (usually shields or other zero-signal circuit lines) connect. *Ground* in an audio system is the primary zero signal reference for the system. In a typical system, there may be a common connection of all audio signal shields and a separate common connection of all power-supply negative terminals. At some physical point in the system, these commons are connected together. That point becomes the system ground and is called the *zero signal reference potential* because that is the place where a voltmeter reference lead is placed. An *earth connection* is a connection directly to the earth often made via a copper rod driven into moist, salted soil. The system ground is physically connected to the earth at this copper rod. These terms are not fixed in meaning, however, and the term ground. for example, is often used in place of either of the other two terms, or grounding may be used as a general term to describe the practice of external noise reduction. In addition, outside the United States, the term earth is often used in place of the term ground to indicate the system zero signal reference potential.

34.5.1.9.8 *Ground Loops in Unbalanced Systems*

A couple of examples will help explain what ground loops are and how to avoid them (also see Chapter 32). In Figs. 34-66, 34-67, and 34-68, the loop is between two audio cables that connect a line-level device to a power amplifier with unbalanced inputs. These are examples of unavoidable ground loops. The best way to deal with this type of ground loop is to keep the cables as short as possible and bundle them physically as close together as possible (lace or tape them together if the setup will allow it). This reduces the area enclosed by the loop, which will reduce the pickup of external noise. Alternately, use an appropriate transformer to balance the connection and allow a telescoping shield to interrupt the ground loop (see the next section).


Chassis connections
Unbalanced cable
Shield
3-wire ac power cable
Ground loop path, follows shield, through device chassis and ac ground in building
Ground loop plan
3-wire ac power cable
Ac ground (part of main ac system)

A. Mixer or other device with unbalanced output.


Chassis connections
Ground loop path
Shield
Shield
Ground loop path follows shield of one cable to chassis of P2100 back through shield of other cable and through mixer chassis.

B. Mixer or other device with stereo unbalanced outputs.

Figure 34-66. Two possible ground loops in an audio system. Courtesy Yamaha International Corp.

Unfortunately, this only reduces the pickup of magnetically coupled external noise. It does not reduce the pickup of common-impedance coupled external noise. Balanced connections and good grounding practice are the only ways to successfully reduce this common noise problem. See Chapter 32 for a more extensive discussion of these problems.

34.5.1.9.9 *Ground Loops in Balanced Systems and Using the Telescoping Shield Connection*

By using balanced connections between two pieces of audio equipment, the shield at the receiving end can be lifted (disconnected) to interrupt one type of ground loop, Fig. 34-69A. This lifted ground results in what is known as a telescoping shield. Since, in a balanced line, the shield does not carry audio signal, it can be disconnected at one end without interrupting the audio signal (and without disrupting the effectiveness of the shield). Unfortunately, this is not a very practical solution to the problem in a portable audio system because it would require special cables that have the shield disconnected on one end (also see Chapter 32).

In Fig. 34-69B the ground loop occurs between a microphone splitter and the two mixers. Even though

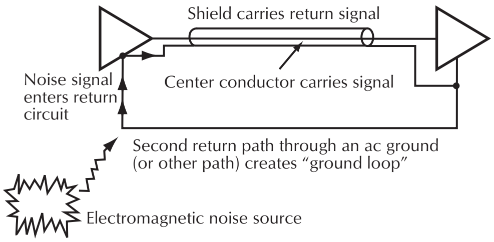
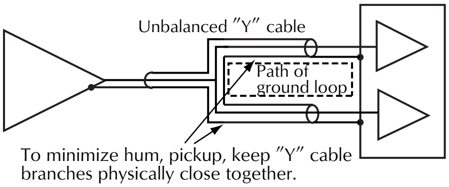
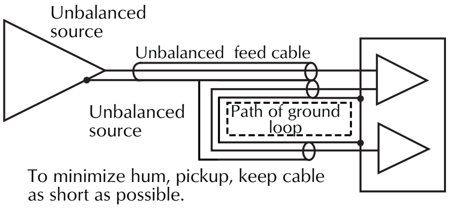

Shield carries return signal
Noise signal enters return circuit
Center conductor carries signal
Second return path through an ac ground (or other path) creates "ground loop"
Electromagnetic noise source

Figure 34-67. Noise entering a system through a ground loop. Courtesy Yamaha International Corp.


Unbalanced "Y" cable
Path of ground loop
To minimize hum, pickup, keep "Y" cable branches physically close together.

A. Unbalanced source device.


Unbalanced source
Unbalanced feed cable
Unbalanced source
Path of ground loop
To minimize hum, pickup, keep cable as short as possible.

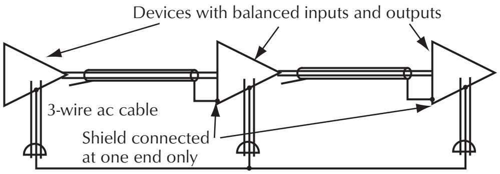
B. Unbalanced source.

Figure 34-68. Minimizing hum with unavoidable ground loops. Courtesy Yamaha International Corp.

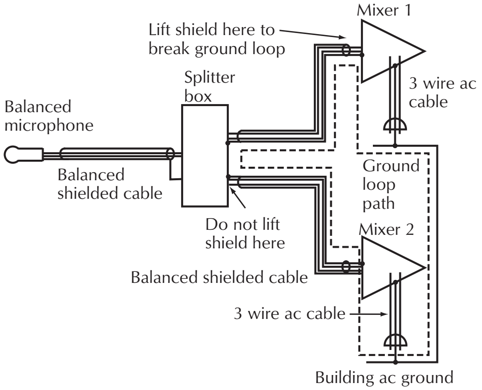
there is only one audio cable connecting the two devices, a second ground connection, through the ac cables of the devices, makes the return connection and forms a ground loop. Using a telescoping shield breaks the ground loop and thus helps prevent pickup of magnetically coupled and common-impedance coupled noise. In Fig. 34-69C a ground-lift switch has been installed so that, when opened, the ground connection through the shield will be interrupted. *Never lift the ac ground on a power amplifier or otherwise defeat the ac safety ground on any piece of equipment.*

34.5.1.9.10 The Pin 1 Problem

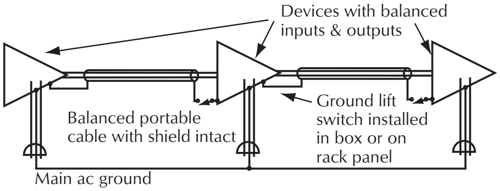
Unfortunately, not all professional audio equipment is properly grounded internally. In some equipment, noise from the shield of a connecting cable is coupled via common impedance into the internal signal path through improper grounding of pin 1 of an XLR connector. Noise caused by this problem will frustrate the best efforts of an experienced installer because the problem is inside the equipment and it cannot be solved externally. The only solution for this problem is to substitute a different piece of equipment with proper internal grounding. Chapter 32 has a more detailed description of this problem.


Devices with balanced inputs and outputs
3-wire ac cable
Shield connected at one end only

A. Telescoping shield (lift the shield at the receiving end).


Mixer 1
Lift shield here to break ground loop
Splitter box
3 wire ac cable
Balanced microphone
Balanced shielded cable
Ground loop path
Do not lift shield here
Mixer 2
Balanced shielded cable
3 wire ac cable
Building ac ground

B. Avoiding a potential ground loop when using two mixers and a microphone splitter.


Devices with balanced inputs & outputs
Ground lift switch installed in box or on rack panel
Balanced portable cable with shield intact
Main ac ground

C. Use of a ground lift switch.

Figure 34-69. Balanced system ground loops. Courtesy Yamaha International Corp.

34.5.1.9.11 Transformers versus Active Balanced Inputs and Outputs

All attempts to solve hum and noise problems can be frustrated if even one piece of system electronics has poor rejection of hum and noise. To avoid this problem, all system electronics should have balanced inputs and outputs (except, of course, for the outputs of system power amplifiers unless they are 70.7 V systems). In addition, the internal circuit design of each piece of audio equipment should be optimized in terms of good grounding and shielding performance.

Both transformer-coupled and active-balanced equipment can provide excellent hum and noise rejection. As a general rule, however, high-quality transformer-coupled inputs and outputs will outperform all but the best active-balanced designs in terms of hum and noise rejection. In addition, transformers offer the benefit of protection against stray dc.

The choice should be made based on a careful examination of the device's specifications, especially the noise rejection performance of the input stages. In addition, remember that lower-quality transformers and simple-circuit, low-performance active balanced inputs and outputs will probably not provide the expected level of hum- and noise-rejection or of audio performance.

34.5.1.9.12 Using Proper Shielding to Reduce Noise Pickup

Proper grounding helps prevent pickup of noise that is transmitted magnetically and noise that is coupled through a common impedance. Magnetically transmitted noise most often comes from motors or, more commonly in audio, from large ac power transformers (either building transformers or the power transformers in a power amplifier or other piece of audio equipment). Proper shielding, on the other hand, helps prevent pickup of noise that is transmitted capacitively. Capacitively transmitted noise may be in the form of radio waves from a radio station or citizens band radio, or it may be in the form of static from certain types of motors or lighting dimmers. Noise from lighting dimmers may also come through the ac lines.

Proper shielding, except in severe noise situations, is straightforward. Use high-quality shielded cables on all microphones and on all line-level equipment, and, if at all possible, install the electronics in a metal equipment rack (preferably steel since this also provides some protection against magnetically coupled noise). Some very low-cost audio cables including guitar cables have poor quality shields. Watch for these potential sources of noise pickup, Fig. 34-70.

It is seldom necessary to use shielded cable for loudspeakers, since they operate at a very high level and a very low impedance. The noise picked up by a loudspeaker cable is actually at the same level as the noise picked up by a microphone cable. However, because the loudspeaker operates at a much higher level than the microphone, the SNR is vastly better, and the noise is seldom a problem.

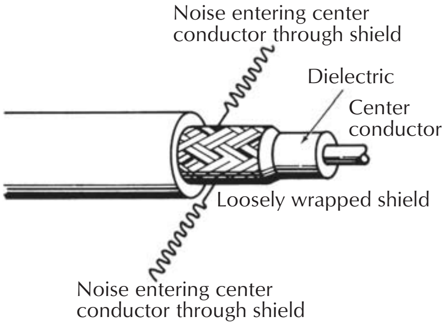

Noise entering center conductor through shield
Dielectric
Center conductor
Loosely wrapped shield
Noise entering center conductor through shield

Figure 34-70. Poor-quality shielded cable. Courtesy Yamaha International Corp.

34.5.1.9.13 Reducing Noise Pickup from ac Lines

Some types of noise, notably noise from lighting dimmers, enter audio equipment from the ac power lines. There are four ways to reduce this problem (see Chapter 32). *A licensed electrician must perform this troubleshooting and any needed modifications.*

1. Install filters on the dimmer circuits (filters at the audio equipment won't help as much and probably will cost a lot more).
2. Make sure the dimmer circuits are properly loaded. In other words, if the dimmers are rated for 1500 W loads, make sure they have 1500 W worth of lighting connected to them. (Or add a suitable dummy load to simulate a full-rated load on the dimmer.) The reason for doing this is that the noise filters (if there are any) will only work properly when the dimmer is loaded properly (this is an example of impedance matching).
3. Be sure the lighting circuits are properly grounded (improper grounding can increase noise levels at the source as well as at the audio equipment).
4. Use a different ac circuit.

34.5.1.9.14 More Tips on Reducing Noise Pickup

Rack Mount the Equipment. Rack mounting, especially when the rack mount rails are made of metal, connects the chassis of all the equipment into a unitized shield. Perhaps more important, rack mounting allows the use of shorter connecting cables and keeps them closer together. When rack mounting large power amplifiers, however, do not place sensitive, low-level equipment right next to them in the rack. The power transformer in a large power amplifier can produce a large alternating magnetic field that can induce hum in low-level equipment.

Keep the Cables Short. Rack mounting can help here, as can simple neatness.

Keep Cables of the Same Type Close Together. Group cables that carry the same signal level. Especially when they form an unavoidable ground loop, keeping cables close together will help reduce noise pickup.

Keep Cables of Different Types as Far Apart as Possible. This means keep the microphone cables away from loudspeaker cables. And keep all audio cables away from the ac power cables. On long cable runs, keep line-level cables and microphone cables separated. It's a common, but risky, procedure to run microphones through a snake (a multimicrophone cable) to a mixer and then run the outputs from the mixer back to the power amplifier through the same snake. This mixing of levels, in a long cable run (greater than about 25 ft) can cause a form of electronic feedback that could cause harmful oscillations in the system mixer.

Keep the Wiring Neat. Carefully made cables, of the proper length (not too long), that are carefully laid out on a stage or in an installation are probably the best way of all to reduce external noise pickup, Fig. 34-71.

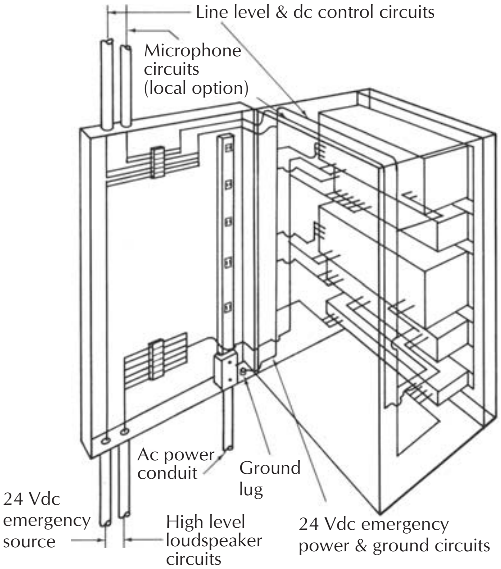

Line level & dc control circuits
Microphone circuits (local option)
Ac power conduit
Ground lug
24 Vdc emergency source
High level loudspeaker circuits
24 Vdc emergency power & ground circuits

Figure 34-71. Cable routing in equipment rack. (After Reference 4.)

34.5.2 Testing and Adjusting

34.5.2.1 Signal Delay in Sound Reinforcement

There are two uses for signal delay in sound reinforcement. The first is to delay one loudspeaker system to allow the sound from a remote loudspeaker system to catch up. This avoids the creation of an artificial echo. The second purpose of signal delay in sound reinforcement is to line up the wave fronts from the components in a packaged loudspeaker system or, similarly, to line up the wave fronts of the various loudspeakers in a loudspeaker cluster.

34.5.2.1.1 Signal Delay for Loudspeaker Clusters

To calculate the delay for a rear cluster in a two-cluster system, choose a typical listener in the coverage pattern of the second cluster who can still hear the first cluster. Calculate the distance from this listener to the first cluster and subtract the distance from this listener to the second cluster. Perform this calculation for several listeners in the coverage of the second cluster who can also hear the first cluster. Choose an average value, biased toward those listeners who can hear the first cluster best. Multiply this average value times 1.13 for distances in feet or 3.71 for distances in meters, to obtain the starting point for delay in milliseconds. Add 6 to 20 ms (the exact amount is best determined on site by listening) to take advantage of the localization known as the Haas effect.

34.5.2.1.2 Signal Delay for an Under-Balcony Distributed System

Choose a listener near the front of the area under the balcony, one who can hear both the central cluster and the under-balcony distributed system. Then, follow the instructions in Section 34.5.2.1.1 above.

34.5.2.1.3 Signal Delay for a Loudspeaker or Loudspeaker Cluster

See Section 34.3.2.11 for a discussion of this topic.

34.5.2.2 Equalization

34.5.2.2.1 The Concept of Equalization

Sound system equalization is a process of adjusting the electronic frequency response of a system to compensate for uneven loudspeaker response and room acoustics. The goals of equalization are to provide a natural-sounding system with good intelligibility and to minimize feedback that might be caused by peaks in the frequency response. In entertainment sound reinforcement systems, equalization may also be used to enhance the sound quality of, for example, a nasal-sounding performer's voice. This use of equalization is very different from the other uses and, in general, it is better to avoid using the same equalizer to both equalize the overall system and provide enhancement of an individual performer.

34.5.2.2.2 What Equalization Can and Cannot Do

Equalization can make a well-designed system sound subjectively better. It can improve intelligibility by smoothing the frequency response or by deliberately peaking response in the intelligibility frequencies (approximately 1500–5000 Hz). Equalization can also help minimize feedback caused by frequency-response irregularities.

Equalization cannot make a poorly designed system sound good. Equalization cannot significantly improve the sound quality of poor-quality loudspeakers. Equalization cannot affect the reverberation time in a room in any way.

Equalization cannot significantly improve a feedback problem in a room where the PAG (potential acoustic gain) is unacceptable. And equalization cannot solve system response problems when those problems are caused by signal alignment irregularities.

34.5.2.2.3 System Design Criteria for Equalization

Equalization should be the icing on the cake for a well-designed system. A system to be equalized should be designed using high-quality, low-distortion loudspeakers and electronics with adequate head room. In a multiloudspeaker system, solve any signal-alignment problems before attempting to equalize the system (see Section 34.3.2.11).

34.5.2.2.4 System Equipment

The system to be equalized must have an equalizer permanently installed in its signal chain. The equalizer must remain in the system after the equalization process has been performed and thus should be a high-quality device with low distortion, low noise, and balanced input and output. Filters should combine smoothly between sections. For a parametric, this can be accomplished by using as broad a bandwidth as possible during equalization and overlapping the bandwidth of adjacent filters just enough to keep the electrical response smooth. Commonly, the equalizer will be a module in a DSP. Choose the module to meet the above criteria.

34.5.2.2.5 Test Equipment

Test equipment should include a calibrated, flat-response microphone, a ⅓-octave real-time audio spectrum analyzer, Fig.34-72, (often called a real-time), and a pink-noise generator. All equipment should be high quality and properly calibrated. If available, a sound level meter, with a flat response position, is useful. Some meters have a line-level output and can thus be used as the system test microphone, Fig. 34-73. Note that newer, computer-based test instruments, such as SIA Software's SMAART system, Fig. 34-74, offer a real-time analyzer function and may thus be used in the equalization process.

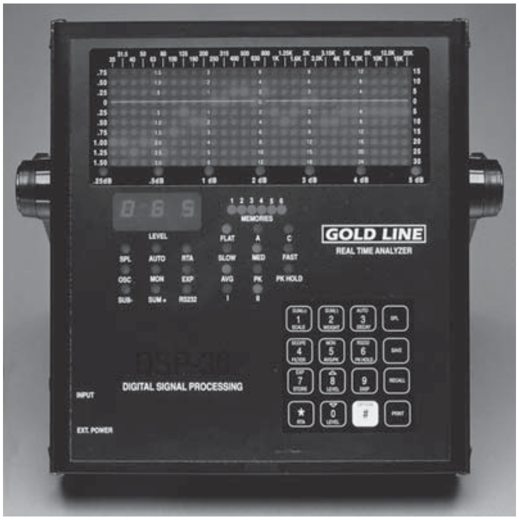


Figure 34-72. A real-time, ⅓ octave, audio spectrum analyzer. Courtesy Gold Line.

Often, a house microphone is substituted for the calibrated microphone in the equalization process. This allows the system response to be adjusted to compensate for the response of the microphone as well as the loudspeakers and the room acoustics. The house microphone, however, should only be used when there is only one type of microphone in the system (and all such microphones have manufacturers' response curves that are very similar to each other).

If more than one type of microphone is used in the system, either a separate equalizer must be used for each type of microphone (indicating a separate mixer) or the calibrated microphone should be used for the

Figure 34-73. Precision integrating sound-level meters and accessory filter sets. Courtesy Brüel & Kjaer Instruments, Inc.

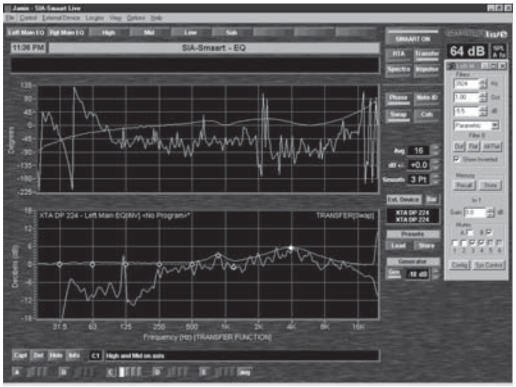


Figure 34-74. SMAART real-time analyzer function. Courtesy EAW/SIA Software.

equalization process and mixer tone controls used to adjust for different microphone types (a more reasonable approach).

34.5.2.2.6 The Test Setup

Connect the pink-noise generator to a typical microphone input (use a pad if necessary to connect a line-level pink-noise generator to a microphone input). Set all system tone controls to their flat positions, and set the equalizer controls to flat. Place the real-time analyzer near the equalizer and place the test microphone in a typical listening position. Avoid placing the microphone directly on-axis of any individual loudspeaker, Fig. 34-75.

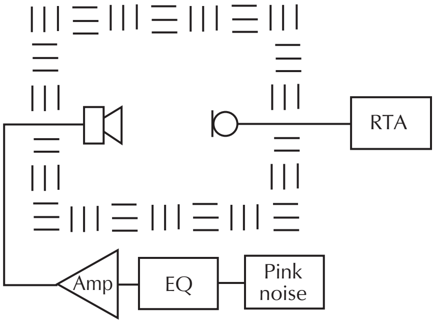


Figure 34-75. Equalization test setup diagram.

Turn on the system and the pink-noise generator, and using the sound-level meter in its flat and slow response positions, set the system gain for the design level. That is, if the system was designed for 90 dB with 10 dB of head room, increase the system gain until the meter reads 90 dB. If the system produces the proper output, reduce the signal at least 10 dB so the system will not go into clipping in any octave.

Turn on the real-time analyzer and observe the response. Note any significant peaks or dips. Move the test microphone to several different locations and note the changes. If the analyzer has memories, these can be used in comparing microphone positions. If the system response changes radically from position to position or has significant peaks or dips at any position, attend to these problems before beginning the equalization process.

34.5.2.2.7 The Equalization Process

To equalize the system, first adjust the high-frequency and low-frequency power amplifiers (in a biamplified system) for the flattest response as indicated on the real-time analyzer. Begin the process of adjusting the equalizer by observing the real-time analyzer and choosing a frequency area that peaks above the rest of the response curve. Using the equalizer, reduce the response in this frequency area. In the beginning of the process, avoid cutting or boosting any frequency more than about 3 dB since later adjustments of adjacent filters may affect the earlier adjustment. Do the same at any other significant peaking areas. If the system equalizer includes boost capabilities, boost carefully between system peaks if you desire, being sure that you are not trying to boost a diaphragmatic absorber (again, no more than about 3 dB at first) to help smooth system response. At this point, a smooth, not flat, response is the goal.

In many two-way, central-cluster systems, response peaks will come in the neighborhood of 200–2000 Hz. These frequencies include the efficiency peaks of the low- and high-frequency loudspeakers and reflect the normal midrange deficiency of many two-way systems. These peaks are, therefore, the first place to make system adjustments, again, working for smooth response at first.

34.5.2.2.8 The Desired System Response Curve

Once the system response is reasonably smooth, begin adjustments toward the desired system response curve, as shown in Fig. 34-76. For a speech-only system, the final curve should be relatively flat from its low-frequency limit (about 50–80 Hz) up to about 1000 Hz. At 1000 Hz, the response should begin a roll-off of about 3 dB/octave to about 8–10 kHz. Response above this frequency should roll-off more rapidly (use a system low-pass filter if available). For a music or music and voice system, begin the roll-off at about 2000 Hz and allow the response to follow this roll-off to 12.5 kHz or higher with rapid roll-off above the desired maximum frequency.

This high-frequency roll-off is a guideline and should be modified for each individual system on the basis of subjective sound quality. The purpose of the roll-off is to improve the system sound quality, since a perfectly flat system response, as displayed on the real-time analyzer, will sound overly crisp, and vocal sibilants (high-frequency breath sounds) will be overly emphasized.

The final equalization curve will probably not follow these rolloff curves perfectly. A final system curve within ±2 dB of the desired curve is, in most cases, more than adequate. Avoid filter settings more than a 6 dB boost or cut whenever possible, remembering that a 6 dB boost requires four times the amplifier power output and four times the loudspeaker power capacity. For this reason, a final equalization curve that requires only a ±3 dB setting on any individual filter but is within only ±3 dB of the desired curve may sound better than a final curve that is within ±1 dB of the desired curve but required ±6 dB of equalization at several filter positions (and, therefore, reduced system head room and increased system noise).

34.5.2.2.9 Why High-Frequency Roll-Off Is Required in Equalization

The reverberant field in most rooms is dominated by the low frequencies. When the equalization microphone is placed in the reverberant field (past the D_c, critical distance from the cluster), the frequency response shown on a real-time analyzer will also be dominated by the low frequencies. From another point of view, the low frequencies don't attenuate much past D_c because they are supported by the reverberant field. The high frequencies, however, continue to attenuate past D_c since they are not supported as strongly by the reverberant field. For this reason, the real-time analyzer will show a frequency response that is dominated by the low frequencies (bass heavy).

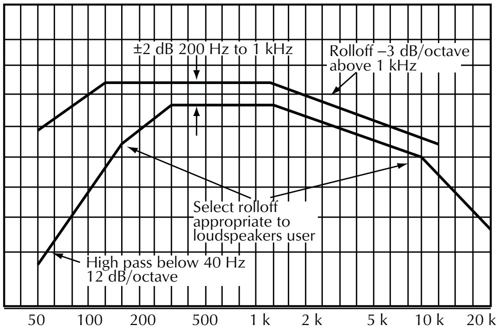
If equalization is then performed using this real-time analyzer display, which is low-frequency dominated, most people will tend to boost the high-frequency response of the system (or cut the low-frequency response) to make the analyzer display a more flat response. This, however, causes the direct sound from the loudspeaker system to be dominated by the high frequencies and our perception of the sound quality will then be that it is too sibilant or too harsh or lacking in bass. The rolloff curves shown in Fig. 34-76 were developed to avoid this problem.

The experience of those engineers who have experimented with direct-sound-only equalization using TEF analyzers (not real-time analyzers) supports the idea that we judge the response of a system based on the direct sound and not on the direct plus reverberant sound. The discussion of the need for roll-off in the previous paragraphs also supports this concept.

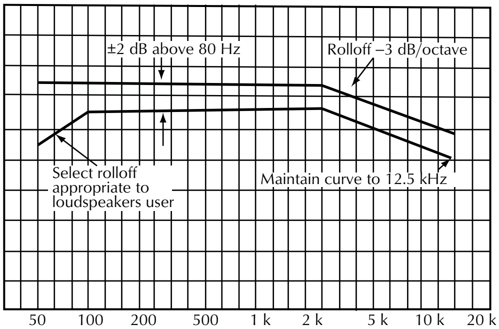
One additional reason to roll off the high-frequency response during equalization is to compensate for the presence boost that exists in many microphones. Listen to the system using the house microphone to see if any additional roll-off is needed for this problem.

34.5.2.2.10 Use of High- and Low-Pass Filters

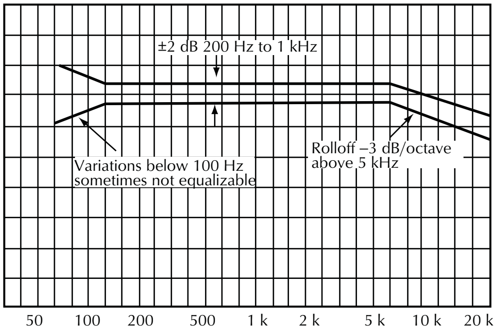
A 12 or 18 dB/octave high-pass filter, at approximately 50–160 Hz, will enhance the performance of a voice-only system by filtering out unwanted low-frequency transients like dropped microphones and breath pops. For a music or music-plus-voice system, use a high-pass filter at 20 Hz or above. Most music systems are actually improved by a high-pass filter at 40–80 Hz. In addition, a vented-box-type, low-frequency enclosure should be high-passed at a frequency slightly below the



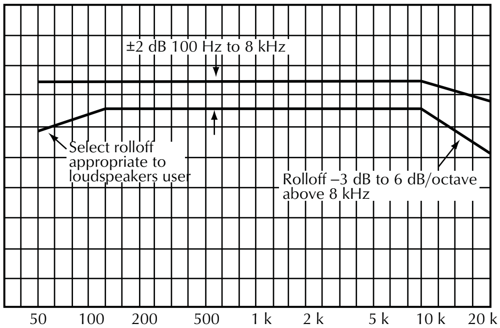
A. Recommended response curve for speech reinforcement systems.



B. International standard response curve for cinema playback systems.



C. Recommended response curve for studio control room monitoring systems.



D. Recommended response curve for high-level rock music reinforcement systems.

Figure 34-76. Various recommended response curves for equalization. Courtesy Bosch/Electro-Voice.

box resonant frequency f_b to help protect the loudspeaker from overexcursion.

A 12 or 18 dB/octave low-pass filter at 12.5–20 kHz will reduce unwanted RF signals and will help prevent system electronic oscillation. In a music system, a low pass at 16–20 kHz will improve the system for the same reasons. Any system with a high-pass filter should also have a low-pass filter for frequency-response balance (the sound quality will be improved).

High-pass and low-pass filters are often included as part of the system equalizer and are available modules in almost every DSP device.

34.5.2.2.11 Equalization Using TEF or Other Equipment That Measures Direct Sound Response

Some users of Time-Energy-Frequency (TEF) test equipment, and other test equipment that can measure direct sound response, have reported very good success in equalizing the direct sound from the loudspeaker system, ignoring (at least temporarily) the frequency response of the reverberant sound. This tends to support the idea that the primary usefulness of equalization may be to smooth the response of the loudspeakers, not the room.

When equalizing the direct sound using TEF-type equipment, do not follow the high-frequency roll-off curves indicated in Fig. 34-76. Instead, equalize for a relatively flat response. Then, begin a gradual high-frequency roll-off at 10–12 kHz. Finally, adjust the overall response for a subjectively pleasing sound while remembering the goals of smooth response and intelligibility.

34.5.2.2.12 Use of Narrow-Band Filters in Equalization

Very narrow-band filters are sometimes added to an equalized system for the purpose of controlling feedback and the ringing that occurs in a system that is near feedback. Used in a well-designed and carefully equalized system, narrow-band filters can be successful for these purposes. Common filter types include parametric filters tuned to a very narrow bandwidth and active and passive narrow-band or notch filters.

Narrow-band techniques for controlling feedback work best in fixed systems in rooms with constant microphone and loudspeaker positions. Even in a portable system, a skilled operator may be able to readjust a set of narrow-band filters to make them useful for feedback control. As in any attempt at feedback control by filtering or equalization, only two or three feedback (or ringing) frequencies can usually be eliminated before a point of diminishing returns is reached.

34.5.2.2.13 Equalization Using Source-Independent Measurement (SIM)

Developed by John Meyer of Meyer Sound Laboratories and Brüel & Kjaer Instruments, the Source Independent Measurement method of equalization uses the sound source, voice or music, as the equalization test signal and an FFT (Fast Fourier Transform) analyzer as the test equipment. While more complex than traditional equalization, the promoted advantages of this equalization method are that it can be used during a concert performance to correct and recorrect changes in system response due to changes in audience size or room humidity.

34.5.2.2.14 Automatic Equalization

Some manufacturers have offered systems that, when properly set up, could actually equalize themselves. However, most designers believe that equalization is complex enough to require the engineering judgments of an experienced human operator. For this reason, there are few such systems on the market, Fig. 34-77.

Figure 34-77. A (discontinued) self-equalizing sound system. Courtesy JBL Professional.

34.5.3 Rigging the Cluster

There are three sections to every rigging system:

1. The loudspeakers and their internal hardware.
2. The building structure.
3. Everything between the loudspeakers and the building structure.

The loudspeaker manufacturer must certify that the loudspeaker is designed for suspension. *Do not suspend any loudspeaker unless the manufacturer has certified it for suspension.* The manufacturer must also specify any limitations on rigging such as the number (or weight) of additional loudspeakers that can be suspended below a single loudspeaker.

For a new facility, the building architect or structural engineer must certify the building structure as being capable of supporting the weight of the cluster with a suitable design factor (safety factor). The system designer must supply the architect or structural engineer with the installed cluster weight. For an existing facility, the owner or system designer should contact the original architect or structural engineer for approval to suspend the cluster in the desired location. If the original architect or structural engineer is no longer available, find another architect or registered professional engineer (PE) to inspect the structure and approve the system suspension in the desired location(s).

The rigging system itself is the responsibility of the system designer and the installing contractor. This includes rigging cables, any suspension grid, and all associated hardware (the loudspeaker manufacturer may supply eyebolts for its loudspeakers). The system designer or installing contractor must present the rigging system design, with associated drawings, to an outside registered architect or professional structural engineer for official approval.

The system must be installed by experienced, professional riggers. Always use hardware that is designed and certified for rigging usage, Fig. 34-78. The system designer should supervise the cluster installation to confirm the aiming points of each loudspeaker.

When finished, the loudspeaker cluster should be inspected for proper loudspeaker aiming and rigging safety. Document the entire installation with as-built drawings.

34.5.4 System Documentation

Here are the most important system documents and their purposes.

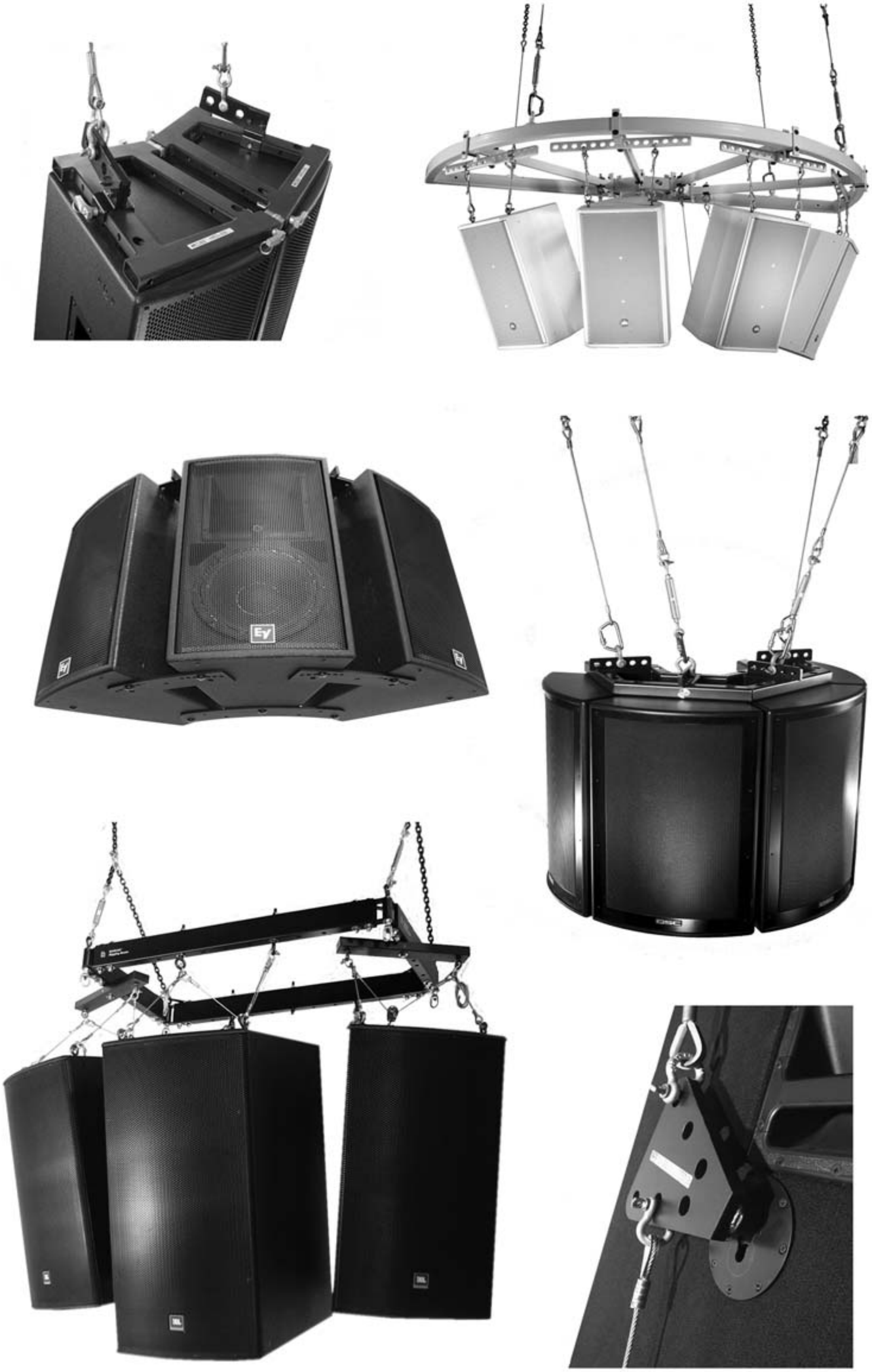


Figure 34-78. Certified rigging hardware. Courtesy ATM Hardware.

34.5.4.1 Original Plans and As-Built Drawings

As-built drawings are critical for service technicians and for a designer of any future system expansion. Keep a set of original plans as reference for any changes that were made. Include a final block diagram (one-line or riser diagram) in the as-builts.

34.5.4.2 Equipment Lists and Equipment Owner's Manuals

Keep an accurate final equipment list. Both the installing contractor and the end user should keep owners' manuals since these may not be available if equipment becomes obsolete.

34.5.4.3 Software Configurations and Backups

Backup the configuration files for DSP systems. Sound systems may be used for a decade or more before being replaced. For this reason, it's a good idea to keep copies of the software used to configure/program the DSP since newer versions may not work on older hardware.

34.5.4.4 Approvals and Certifications

Keep rigging drawings and approvals and any other safety agency approvals such as those from a local fire marshall.

34.5.4.5 Additional Documentation

Keep a written record of all system settings such as amplifier-level control settings and equalizer control settings.

34.5.5 Troubleshooting a Sound System

Repairing a sound system may require the skills of a trained technician. Troubleshooting, that is, finding the problem, is something almost anyone can do if they

1. Know the block diagram of their system.
2. Understand what each component in the system is supposed to do.
3. Know where to look for common trouble spots.

34.5.5.1 Know the Block Diagram

A sound system block diagram explains how the various components in the system are connected to each other and what happens to a signal as it flows through the system. Because the block diagram shows the way the sound system operates, it is extremely useful in the troubleshooting process.

As obvious as it may sound, it's not possible to tell whether a component is working properly or not unless its original function is well understood. Thus, it's a good idea to keep instruction manuals on all components handy. Some repairs are as simple as repositioning a control knob or throwing a switch that someone has inadvertently changed.

Cables and connectors are by far the most common sources of problems in audio systems. This is the best reason to keep lots of spares, especially of cables that are moved around a lot, such as microphone cables.

Other common trouble spots are fuses and circuit breakers, as well as switches and controls that are in the wrong positions, and problems with house ac power.

34.5.5.2 Logical Troubleshooting

The process of troubleshooting involves logical thought and methodical tracking down of a problem by elimination.

Logical thought processes come into play when a problem first occurs. If a single microphone goes suddenly dead, logic says that the power amplifier probably isn't at fault. If, on the other hand, an entire system is suddenly quiet, the power amplifier might be at fault, because it's not likely that all the microphones have failed at once.

A methodical elimination process, as shown in Figs. 34-79 and 34-80, can track down the source of most problems very quickly. The idea is to find out what component (microphone, cable, mixer, amplifier, loudspeaker, etc.) is causing the problem and to replace or repair it. During a live performance, of course, replacing a faulty component is the most likely cure since a repair might take up too much time.

The system mixer is a good place to begin the troubleshooting process because it has the controls for the entire system. A noise in the system, for example, can be traced by looking at the VU meters or listening to a PFL (pre-fade-listen) circuit with headphones. This alone, may indicate that the noise problem is coming from one microphone. Pull down the fader for that input channel. If the noise goes away, check out the microphone, or more likely, the microphone cable.

If the entire system suddenly goes dead, again, check out the VU meters or listen to the PFL. If they are still active, then it is most likely that the system is working through the mixer (it's still possible that the output circuits of the mixer are at fault). Thus, some compo-

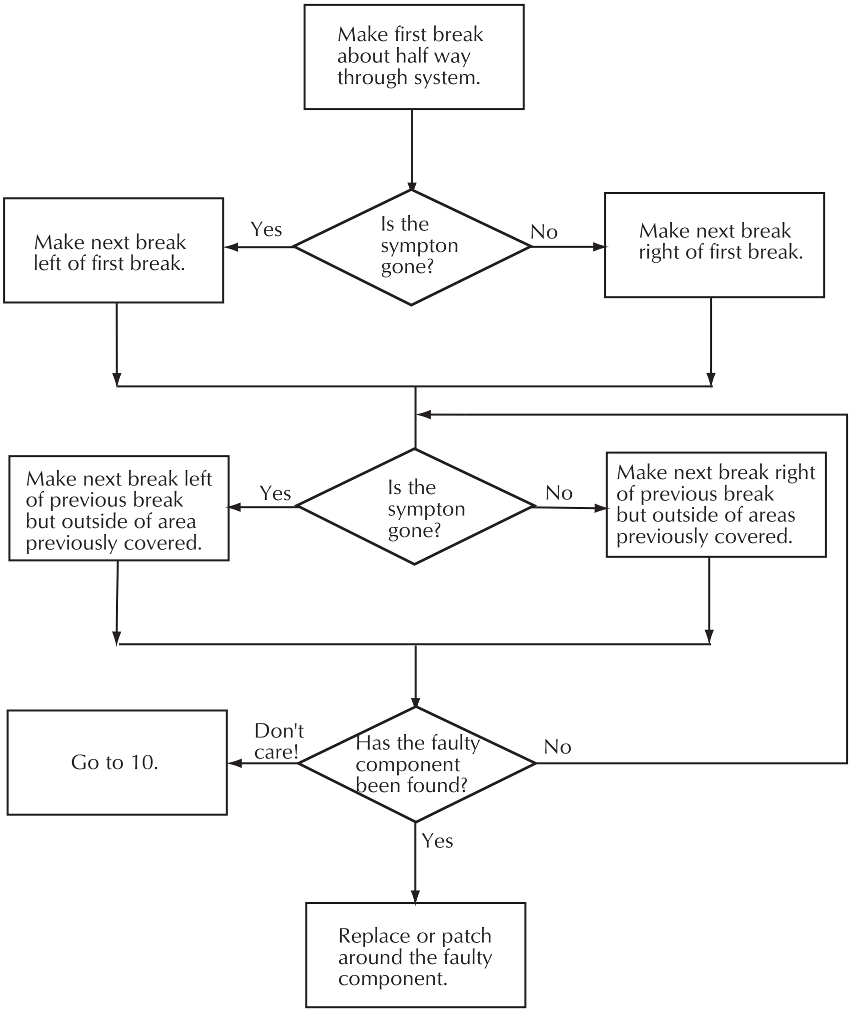

Make first break
about half way
through system.
Is the
sympton
gone?
Yes
No
Make next break
left of first break.
Make next break
right of first break.
Is the
sympton
gone?
Yes
No
Make next break left
of previous break
but outside of area
previously covered.
Make next break right
of previous break
but outside of areas
previously covered.
Has the faulty
component
been found?
Don't
care!
No
Yes
Go to 10.
Replace or patch
around the faulty
component.

Figure 34-79. Troubleshooting Part 1. Assumes that the problem is hum, noise, or oscillation and that block diagram flow is from left to right. Method is to break system (disconnect) at indicated points until the faulty component is located.

nent farther along in the system is the most likely culprit. Think through the block diagram at this point to find the next suspect component. (One component that may be a problem is the house ac power.)

When possible, patch around suspect components. For example, a limiter can be completely removed from the system, and the system will still operate. Thus, if a limiter is suspect, use a patch cable to bypass it. If the bypass operation causes the system to begin operating again, the limiter is at fault.

When the suspect component is necessary to the operation of the system, try to replace it with some other equivalent component. If a loudspeaker is suspect, for example, try switching it with a similar loudspeaker or even a stage monitor loudspeaker temporarily in place of a main system loudspeaker. If the mixer is suspect, try running a CD player or MP3 player directly into the system power amplifier to make sure that portion of the system is still working.

34.5.5.3 DSP Troubleshooting

It's common for an all-in-one DSP unit to host all of a sound system's signal processing—everything between the mixer and the power amplifiers. A suspect all-in-one DSP like this can't be bypassed for troubleshooting. Also, it's not easy to swap it for another DSP, even if one is available, because the second DSP must be pro-

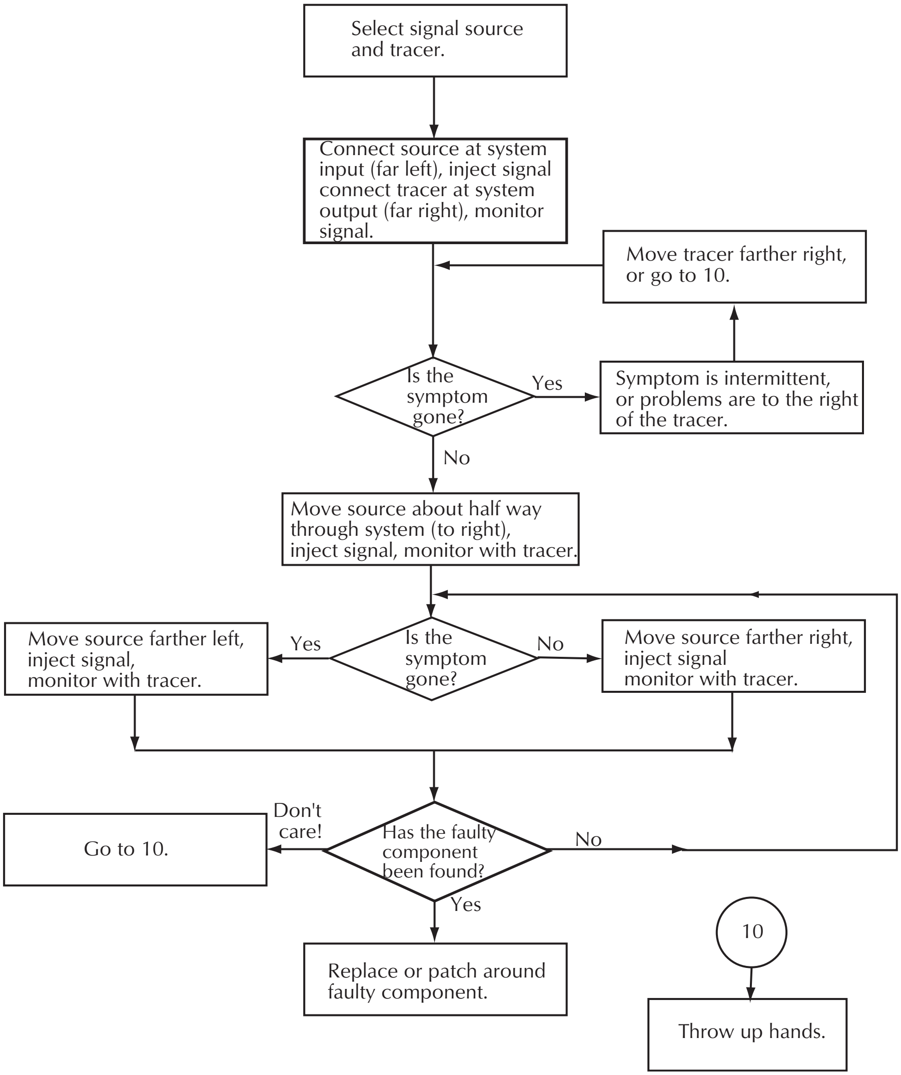

Select signal source and tracer.
Connect source at system input (far left), inject signal connect tracer at system output (far right), monitor signal.
Move tracer farther right, or go to 10.
Is the symptom gone?
Yes
Symptom is intermittent, or problems are to the right of the tracer.
No
Move source about half way through system (to right), inject signal, monitor with tracer.
Move source farther left, inject signal, monitor with tracer.
Yes
Is the symptom gone?
No
Move source farther right, inject signal monitor with tracer.
Go to 10.
Don't care!
Has the faulty component been found?
No
Yes
10
Replace or patch around faulty component.
Throw up hands.

Figure 34-80. Troubleshooting Part 2. Assumes that the problem is distortion or interruption of signal and the block diagram flow is from left to right. Method is use of signal generator and tracer (a tape deck and powered loudspeaker may suffice) to locate the faulty component.

grammed and adjusted exactly the same as the first DSP. If you suspect an all-in-one DSP, try patching from the mixer directly into the system power amplifiers. Keep the level down on the mixer and insert the mixer's input-channel high-pass filters to avoid damaging HF loudspeakers. This is a crude method but should provide some insight into the problem.

A second problem with DSP troubleshooting is that the trouble may be improper DSP programming or adjustment. For example, if a graphic equalizer module has several adjacent filters reduced by several dB, it may be necessary to increase the gain in the following module to compensate. Then, if a high-level signal occurs that's outside the frequency range of those filters, it may overdrive the following gain stage. It's not easy to troubleshoot this kind of problem because you can't insert real test equipment between the DSP modules. However, many DSP units now include meters and signal generator modules that can be patched into system nodes for troubleshooting.

34.5.5.4 Anticipating and Preparing for Problems

These techniques can help troubleshoot DSP problems and other problems but this kind of troubleshooting can take time. Thus, it's better to anticipate potential problems in the system and prepare to solve them quickly. Keep spare cables, spare components, and troubleshooting equipment nearby. Plan what to do if a critical system component fails. This includes the system mixing console and any all-in-one DSP units. Some facilities with critical sound requirements keep a smaller mixer and a second, programmed DSP ready to swap in case of failure. Finally, when possible design the system to minimize problems resulting from the failure of a single piece of equipment. For example, replace the all-in-one DSP with networked amplifier DSP units. This way, the failure of a single DSP cannot cause the entire sound system to fail.

34.6 Applications

34.6.1 Portable and Tour Sound Systems

Portable systems range from voice-only paging systems, as might be used at a local county fair, up to the giant tour sound systems used for outdoor rock music festivals. The design criteria for a portable system build on the criteria discussed for permanently installed systems, adding to and modifying the system to take into account such obvious considerations as travel and less obvious considerations such as the potential for abuse by an inexperienced operator or an overexcited audience.

34.6.1.1 Packaging

Portable systems must be rugged to survive travel. They must be packaged efficiently in order to fit into as small a travel vehicle as possible and so that setup and teardown are quick and efficient. Even large tour sound systems are normally designed to fit efficiently into standard-size trucks (48 foot or 52 foot semitrailers in the United States). Efficient packaging leads to lower ownership and operating costs from the ability to use smaller vehicles (or fewer large vehicles) and from the requirement of fewer hours for setup and teardown. Rugged packaging for all components of the system, from loudspeaker systems to electronics to microphones and accessories, also reduces system maintenance costs and improves reliability.

34.6.1.2 Loudspeaker Systems

In the past, a county fair might rent a portable system consisting of a group of 70 V paging horns, matching electronics, and one or more paging microphones. Today, most portable systems use packaged loudspeakers or line arrays and the largest tour sound systems may even use proprietary loudspeaker systems designed and built by the tour sound company itself.

For efficient arraying, most portable packaged loudspeaker systems are trapezoidal in shape. Manufactured packaged loudspeaker systems are usually two-way or three-way designs. Proprietary, tour-sound packaged loudspeaker systems are commonly three-way or even four-way systems.

Although some four-way proprietary tour-sound systems include subwoofers, most popular music applications use separate subwoofers. Separate subwoofers mean the main packaged loudspeaker systems can be smaller. Putting subwoofers on the floor reduces the size of any suspended array.

Smaller portable systems may use a single type of packaged loudspeaker. For example a weekband band may use a two-way or three-way, 90° design and find this fills its needs for nightclubs and smaller dance halls. Larger systems will benefit from more than one type of packaged loudspeaker system. For example, a mid-sized tour sound system might include both 60° and 90° packaged loudspeaker systems along with separate subwoofers.

In the past, tour sound systems used separate components (horns and woofers) in place of today's packaged loudspeaker systems. While systems using separate components took longer to set up and tear down, they provided increased versatility because the operator could choose from multiple horn patterns and design a custom array for each venue.

A designer of a portable system today could bring back some of this versatility by including both 60° and 90° packaged loudspeaker systems, by using separate subwoofers and by packaging a few component long-throw (40° by 20°) horns to cover the upper seats in an arena or the far rows of an outdoor theater.

Many tour sound companies are now using line array-type systems. These systems utilize specially designed loudspeaker systems and sophisticated DSP electronics to create precisely controlled vertical directivity patterns that can reach the back of an audience with good sound quality while maintaining a reasonable level in the front.

Selected models of line arrays and packaged loudspeaker systems are available in self-powered versions with internal amplifiers and often with internal DSP signal processing, which may include crossover, delay, limiting, and equalization. If designed for portable use (and suspension), these systems can eliminate the need to carry racks of amplifiers and signal-processing gear. However, they complicate suspended system design because ac cabling must be included in the array.

34.6.1.3 Portable System Rigging

Rigging a portable system is much like rigging a permanently installed system (see Section 34.5.3). Do not suspend any loudspeaker system unless its manufacturer has certified it for suspension. Get the entire rigging system approved by a licensed architect or professional engineer (PE). For each new venue, consult with the building architect or a local PE to confirm that the building structure is capable of supporting the system with adequate design factor (safety factor). Have the system rigging performed by certified rigging professionals.

34.6.1.4 Cabling

The cabling system for a portable system must be every bit as rugged and efficient as the packaging for the loudspeaker systems. Solid-core wire is absolutely out of consideration. High-quality, multistrand wire of a large wire size (low gauge number) is highly recommended.

Loudspeaker connectors must be rugged, high capacity, and easy to connect (but difficult to connect improperly!). While smaller portable systems may utilize high-current phone plugs, most larger systems use a specialty type of twist-lock loudspeaker connector, manufactured by Neutrik and known as the Speakon. Some amplifiers even have Speakon connections.

Microphone and line-level signals require equally rugged cable with the addition of a high-quality shield. Wire must be stranded, and the shield should be a tight braid of stranded wire. Foil shields will crack and break from the continuous flexing of portable use. Microphone cables should have a highly flexible outer sheath. One exception to this rule is the snake cable, a group of individually shielded, twisted-pair cables in one outer sheath. The shields in a snake cable are foil to reduce the overall diameter of the cable, and the outer sheath is usually made of vinyl or some other plastic-type material that is less flexible than the rubber sheath used on microphone cables. These compromises demand special care for the snake cable, especially considering its high cost per foot. The cable must be coiled carefully in storage and, when it is in use, a mat or other protection must be placed over the cable in any area where it might be walked on.

It's a common, but risky, practice to run microphone signals from a stage area to a mixing area through the same snake cable used to feed line-level signals back to the stage electronics. The snake provides a transformer-like coupling between the inputs and outputs of the mixer. This can turn the mixer into a high-frequency oscillator. The reason some popular music systems seem to operate successfully this way is that the microphone output levels are so high that the mixer gain is low enough to prevent the oscillation. One low-level microphone (like an acoustic guitar microphone), however, is enough to cause the problem. Thus, separate microphone-level and line-level snakes are recommended. Some cable manufacturers are now offering snakes that are specifically designed to reduce the problems caused by running microphone and line-level signals in the same snake. These may be useful in portable systems.

Microphone connectors are invariably XLR type for good reasons. The XLR connector is rugged yet easy to connect and disconnect. It has ample current capacity for microphone and line-level signals, and it has limited self-wiping (cleaning) of its contacts. Furthermore, pin 1 of the XLR always connects first. This allows the shields of the cables being connected to equalize their static charge before the signal wires are connected and thus helps avoid electrical (static discharge) transients. Snakes with only a few cables are often terminated with a group of individual XLR connectors. Larger snake cables often have elaborate multipin connectors that lead to either a group of XLRs or a metal box with a group of chassis-mounted XLRs. Some tour companies have adapted their mixers to accept a multipin snake cable directly. This simplifies connections but eliminates the ability to repatch the cables (at least at the mixer) for a different stage setup, and it hinders quick troubleshooting.

Digital cabling for sound systems borrows from computer network standards. Unfortunately, computer networking connectors such as RJ45 (Ethernet) and fiber-optic connectors, are not rugged enough for constant portable usage. Some manufacturers now offer ruggedized versions of these connectors, Fig. 34-56, and more such connectors will certainly appear in the future.

34.6.1.5 Electronics

Electronics for portable systems are generally the same as for installed systems. Choose the electronics for ruggedness, however, in addition to performance. The large power transformer on a power amplifier, for example, must be securely fastened to its chassis to avoid physical destruction of the power amplifier during the jolts of traveling over rough roads.

Most portable equipment will be rack mounted. Thus, the equipment must be designed to survive the jolts of travel when mounted in a rack. Experienced tour companies often support the rear of large rack-mounted components to help prevent damage. The racks themselves, of course, must be rugged and travel well, and small racks can mean heat buildup so that extra cooling fans should be considered.

Some manufacturers offer electronic packages specifically designed for traveling. The powered mixer, designed for smaller portable systems, is a good example. Some powered mixers include a full-function mixing console, internal effects, one or more graphic equalizers, compressor/limiters, and one to four power amplifiers. For a small- to medium-sized portable system, these powered mixers are often the only electronics needed.

Mixers and other nonrack-mounted equipment must be carried in padded road cases. Similar cases can be used for microphones, cabling, and system accessories.

34.6.1.6 Multichannel Portable Systems

The primary problem with a multichannel portable sound system is that, ideally, each member of the audience should be able to hear all of the individual channels. That means, ideally, each loudspeaker channel must cover the entire audience. The large system required to make this happen prevents true stereo sound in most portable systems. Stereo-type effects, however, can be achieved with traditional left and right loudspeaker systems, and side and rear loudspeaker systems can be successfully used for fill and special effects.

34.6.1.7 Electrical and SPL Safety in Portable Systems

Electrical safety in portable systems is complicated by the uncertain condition of the ac power system of each building. One excellent way to bypass this problem simply is to carry a portable ac power distribution system, which should be designed and constructed by a qualified, licensed electrician. Have a local licensed electrician connect this portable system to the house ac power. Often, the portable ac system can be connected directly to the building ac service entrance (a local qualified licensed electrician must perform this connection). This not only bypasses any potential safety problems in the house ac system but also provides a (relatively) clean ac power system for the noise-sensitive sound system electronics.

High L_P (high sound-pressure-level) hazards are often overlooked but are, nonetheless, dangerous. It is a well-established fact that high L_P, over an extended period of time, can cause permanent hearing damage. Hearing protection is a must, especially when performing high L_P equalization or other testing or when checking out individual loudspeakers before or during a performance. It is possible to wear concealed ear plugs (the expanding foam type are comfortable and very effective) during a concert performance, even if you are mixing the performance. The human brain adjusts the hearing mechanism to the point where things begin to sound right again after a short period of wearing hearing protection. The situation is similar to wearing sunglasses. After a period, colors begin to look right again. This type of hearing protection is important especially for the extremely high L_P encountered during stage monitor mixing. When in doubt about whether or not you need hearing protection, listen to the ringing in your ears after a performance. This ringing, known as tinnitus, is the human body's way of telling us that the sound level is too high. Prolonged exposure to these levels, of course, will almost certainly lead to permanent hearing loss.

34.6.1.8 Performance Criteria

Any system for entertainment must be designed to accept and reinforce an L_P of 100–120 dB. These levels represent amplifier power output and loudspeaker power handling in the neighborhood of 1000 times that of a speech-only system. Microphone input levels are such that electrical output from a high-sensitivity microphone may be as high as 0 dBu on peaks. Thus, mixers must have high-input capabilities and lots of head room.

Ideally, system head room should be as high as 20 dB, although, economically, this may be unobtainable. Extensive use of compressor/limiters can make a 10 dB head room system sound almost as good as one with 20 dB of head room (and the loudspeaker and power amplifier costs are lower).

The frequency response of a system designed for popular music must extend down to at least 40 Hz (or lower) and at least as high as 12–16 kHz. One way to approach the design of this type of system is to design a vocal system with response from about 80 Hz to about 8 kHz and add supertweeters and subwoofers for the very low and very high frequencies. The subwoofers and supertweeters can be considered special effects for this type of system.

34.6.1.9 Stage Monitor Systems

Stage monitors are as important as the house loudspeakers for an entertainment sound system for the simple reason that they aid the entertainers and thus help encourage a better performance.

Stage monitor loudspeakers must be unobtrusive yet high performance. One way to accomplish this is to keep the low frequencies out of the stage-monitoring system so that enclosure size can be minimized. Another way is to place full-range stage monitors on the sides of the stage and limited-range monitors at the performer locations.

Whenever possible, treat the stage monitor system as a completely separate system. Split microphone lines on stage, and send one signal to the house mixer and another to a separate monitor mixer. Use a splitter transformer if possible so that grounding isolation can be maintained. Mix, equalize, and power the monitors separately, with the monitor mixer somewhere in the vicinity of the performers, perhaps at one side of the stage so that the operator can hear the results of mixing actions.

Multiple outputs are needed and useful on a monitor mixing system, since each performer may want his or her own mix. For this reason, several manufacturers offer mixers specially designed for the task of monitor mixing.

Equalization of a monitor mixing system is primarily for the purpose of avoiding feedback. Thus, multiband parametric and notch-type filters may be superior to the more common graphic equalizers.

Newer wireless in-the-ear monitoring systems have become a popular alternative to stage monitor loudspeakers. These systems, which operate on wireless microphone frequencies, allow each performer to receive a custom monitor mix and eliminate the need for monitor loudspeakers on the stage.

34.6.1.10 The Entertainment System as a Musical Instrument

Consider that many of the individual instruments used in modern music cannot exist apart from their electronics and loudspeakers. Add a sound system with multiple loudspeakers and, most likely, multiple phasing problems. Close mic the vocals to pick up breath noises not heard in normal conversation. Choose a microphone with lots of proximity effect so that performers can change the quality of their voices by the way they hold the mic. Close mic even the acoustic musical instruments so that feedback can be avoided but so that normal acoustic mixing of the complex acoustic sources in a musical instrument is eliminated. Mix the signals in a way that has little in common with the acoustic mixing that comes from the geographic layout of an orchestra. Add artificial equalization, reverberation, recorded segments, purposeful harmonic distortion, and other special effects. The result, when the sound system operator is as good an artist as the stage performers, is popular music! And, again, that popular music could not exist without the electronics, including the sound system. Finally, more and more, the sound system used in entertainment bears little resemblance to and cannot rightfully be called a sound reinforcement system. It certainly reinforces, but it also enhances, and, to a very great extent, it creates. Thus, there is ample justification for considering the entertainment sound system to be a musical instrument in its own right.

The significance of this, for traditional sound system designers, is that many of the rules of good sound system design can, and in fact, must, be modified for the design of an entertainment sound system. Perhaps more significant, a very well-designed and operated entertainment sound system can be extremely effective in performing its design goal, that of helping a group of artists to entertain an audience. A nontechnical member of that audience may believe that the particular type of sound system would be the answer to some sound reinforcement problem at a local facility. Technical and nontechnical people, then, should understand that the entertainment sound system is, in actuality, a musical instrument, designed for modern popular music, and only for that purpose. Place the same system in another facility and use it for reinforcement of speech or classical music, and it may be entirely unsuitable.

34.6.2 Systems for Religious Facilities

The primary challenges in designing a sound system for a religious facility are the user interface (many users will be volunteers who are unfamiliar with sound system operation), the aesthetic design (most religious facilities will want the loudspeakers hidden), and the acoustic environment (many religious facilities are, like the typical cathedral. highly reverberant).

34.6.2.1 User Interface

Many users in a religious facility are nontechnical and unfamiliar with sound system operation outside of, perhaps, a home theater system. There are two exceptions: the facility lucky enough to have a trained operator on staff and those facilities that have services that include popular religious music, dramatic presentations, and so on, and make full use of the capabilities of an entertainment-type sound system. These exceptions should receive a system designed with a trained operator in mind.

34.6.2.2 Systems for Inexperienced Operators

It's a good idea to place all seldom used controls and adjustments (including the system equalizer) behind the locked door of an equipment rack or within the software of a DSP system. Minimize the number of controls seen by the user and minimize the complexity of their function. A typical set of controls might consist of a simple mixer with one set of treble and bass controls and a master volume control.

An automatic mixer can be extremely useful since it takes over most of the caretaker functions of mixing. A system with an automatic mixer may need only one user adjustment—the on-off switch. A compressor/limiter and a feedback detector/gain-reduction device would be valuable additions to such a system.

In any such system, provide needed and useful user interfaces such as an MP3 player, CD, or DVD player, and a volume control for each. It may be possible to integrate these into the automatic mixer so that volume controls on the external devices are all that are needed.

Also keep in mind the possibility that an experienced operator may volunteer, in which case the facility may wish to upgrade to a more complex user interface, probably in the form of a mixing console.

A digital mixing console may seem like an ideal way to provide increased capabilities to a religious facility with inexperienced operators. The designer can set up the console with recallable scenes and the operator simply selects the appropriate scene for each worship service. However, the smallest change in a service may necessitate significant changes in settings that require human judgment. When that happens, the digital mixing console may be more difficult to understand than a conventional analog mixing console. Thus, digital mixing consoles may be best for facilities with experienced operators.

34.6.2.3 System Aesthetics

Many religious organizations have large, architecturally beautiful facilities. In general, they do not wish to alter the architectural lines (which may have religious significance) by adding large loudspeakers. In particular, the central cluster type of loudspeaker system almost always ends up in an aesthetically undesirable location.

Religious organizations and their architects should be encouraged to design new buildings with a sound system in mind and to ask a qualified acoustical consultant to join in the process in the early planning stage. Systems for existing facilities, however, must contend with existing architecture and sight lines.

When a central cluster is the right choice for good coverage, it may be enclosed in a framework covered with grille cloth chosen to match the room decor. In those facilities with an attic space above the auditorium it may be possible to hide the system behind a large (new) opening in the ceiling, again, covered with grille cloth. Enclose the cluster above the ceiling, too, to prevent the loss of valuable heat through the hole in the ceiling. Brace the enclosure and line it with fiberglass or other sound absorbent material to help reduce acoustic problems.

Newer Christian churches in the United States often choose a wide, fan-shaped auditorium. For these churches, an exploded cluster is a good way to cover the space with several, relatively small clusters that are easier to disguise. Note that column-style line arrays on each side are usually not a good choice for this style of space because they cannot adequately cover the center of the seating area near the stage.

In a rectangular room, a small central cluster, augmented by a second cluster on delay, may help solve sight-line problems. Column-style line arrays, placed left and right at the sides of the platform, are another alternative for this type of room, providing good coverage while maintaining a discrete profile.

In deep rectangular rooms, a distributed system may be the best solution. In high-ceiling rooms, a distributed system may consist of column-style line arrays or small packaged loudspeakers installed on building pillars. A small localizer loudspeaker in the front can help maintain a natural directionality.

One oft-suggested approach, placing the loudspeakers in an organ chamber, far to the rear of the system microphones and behind a wooden grille, is almost always unworkable. The position of the loudspeakers aims them directly at the system microphones, which usually results in feedback problems. In addition, the wood grille work, which may be fine for organ music, can cause severe acoustic problems (both from cancellations and from vibrations) in the sound system loudspeakers.

34.6.2.4 Reverberant Room Problem

It is said that the types of chanting services employed by some religious groups developed in large reverberant cathedrals before sound systems were invented. Chanting helped carry an intelligible message to the congregation. All too many religious organizations still have the problem of reverberation, yet many have given up the chanting type of service and now want intelligible speech in their facilities.

The problem of intelligible speech in a reverberant religious facility is no different than in any other reverberant room, except, perhaps, for the desire to hide the loudspeakers. One other consideration, however, is that pipe organs and much religious choir music depend on high levels of reverberation, a criteria that conflicts with the desire for lower reverberation times for speech. Because even a good compromise may be expensive, the services of a qualified consultant are invaluable when this situation arises.

There is no magic way to design a loudspeaker system for a reverberant room. However, there are two ways to maximize the intelligibility of a sound system in a reverberant room. First, get the loudspeakers as close to the listeners as possible. This maximizes the direct-to-reverberant ratio at the listener's ears, which improves intelligibility. A distributed system (ceiling type, on building pillars or pew-back) is the easiest way to accomplish this goal. Second, use directional loudspeakers and point them at the listeners and away from walls, ceiling, and other hard surfaces. Sometimes, these two concepts can be combined. For example, in a distributed system on building pillars, use a line-array-type column loudspeaker to direct sound at the listeners.

34.6.3 Sports Stadiums and Other Outdoor Systems

The discussion of outdoor system design presented in Section 34.2.2 was limited to very basic, theoretical considerations. The following discussions include many of the practical aspects of the design of an outdoor system. Also see Chapter 7 for a thorough treatment of sound systems in stadiums and other outdoor venues.

34.6.3.1 Excess Attenuation of High Frequencies in Air

The friction of air molecules rubbing against each other causes attenuation of sound that adds to the loss caused by the inverse-square law. This frictional loss is normally insignificant in indoor systems (except in very large rooms) but can become a problem in large outdoor systems because of the long distances involved. The problem is considerably worse at the high frequencies, which are important for speech intelligibility. This is because the molecules of air are moving faster than at low frequencies. The problem also increases at lower relative humidity as shown in Fig. 34-81.

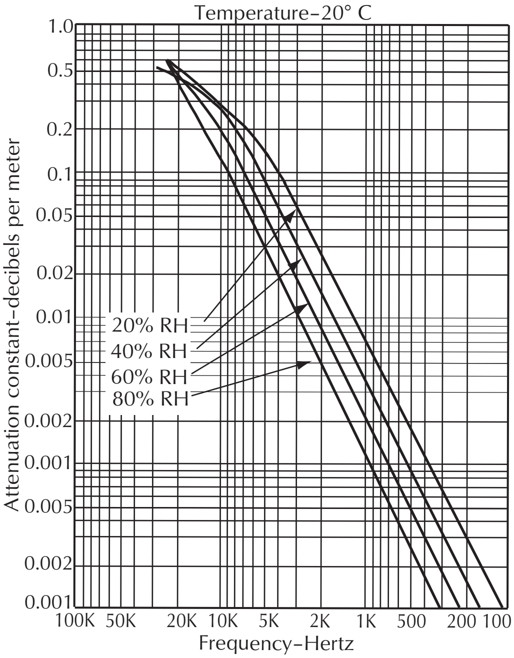


Figure 34-81. Absorption of high frequencies in air. Courtesy Bosch/Electro-Voice.

The problem shows up in outdoor systems with long distances between loudspeakers and listeners. It often cannot be solved with simple equalization or even by adding additional high-frequency horns. The reason is that the attenuation may be 10 to 20 dB or even more, depending on frequency, distance, and relative humidity.

One potential solution is to add additional loudspeakers (or high-frequency horns) at a position nearer to the listeners and, of course, to place these loudspeakers on delay. Since the attenuation of low-frequency information is much less, it is normally unnecessary to add additional low-frequency loudspeakers just to overcome the loss caused by friction in the air.

In many speech-only systems, the loss is simply tolerated. Intelligible, if somewhat telephone-quality speech, does not require frequencies much above about 3 kHz. Except in extreme situations, adding a few extra high-frequency horns and performing some additional equalization will result in an acceptable system. Obviously, the additional equalization should be performed on the long-throw loudspeakers only since the distance from the short-throw loudspeakers to their listeners will be considerably less than for the long-throw loudspeakers.

34.6.3.2 Using Distributed Systems Outdoors

There are several reasons to consider a distributed system for an outdoor facility. Overcoming excess attenuation of high frequencies is one reason. Avoiding the effects of wind and temperature layers is another reason. Distributed systems may reduce annoying neighborhood leakage.

In some stadiums, there is no desirable location for a central cluster. A round or oval stadium with full round seating is one example. The usual scoreboard location will often be awkward for coverage of nearby seating. Sometimes existing lighting blocks the only workable cluster location. In these or other similar situations, a distributed cluster type of system may be the solution. The clusters can be placed under balconies (under one seating section to cover the one below) or on lighting poles. It is acceptable to place loudspeakers behind the heads of the spectators if this location does not cause artificial echoes or other problems.

The distributed approach may also work well in small outdoor systems such as a high-school football field system. When the audience sits in one or two relatively small sets of bleachers on either side of the field, it is often easier to place horns on existing lighting poles near the bleachers than to build a large central cluster at one end of the field, Fig.34-82.

In any distributed system, consider sight lines and watch out for potential artificial echoes.

34.6.3.3 Echoes Outdoors, Artificial and Otherwise

Normally, echoes are created by sound from a source, such as a loudspeaker cluster, reflecting off a hard surface and reaching a listener at a time that is delayed enough to be perceived as an echo. About the only way to deal with these echoes, outside of treating (or remov-

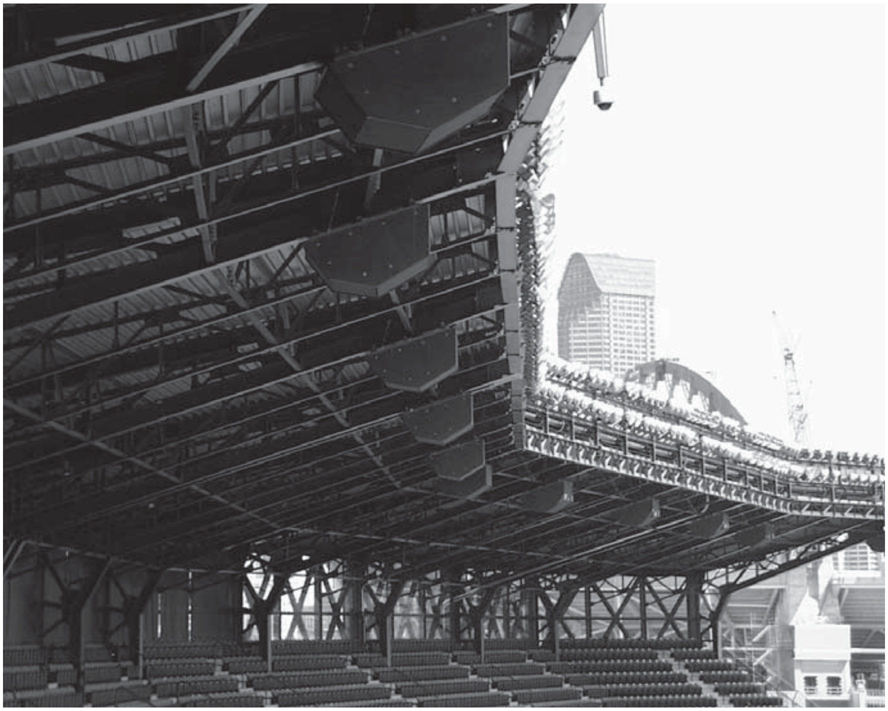


Figure 34-82. Distributed clusters in a stadium. Courtesy Seatle Mariners.

ing) the reflecting surface, is to simply aim the loudspeakers away from the offending surface. Narrow coverage angle loudspeakers and constant-directivity horns may help.

Echoes may also be created by the sound system because of poor layout or because of poor use or nonuse of electronic signal delay. Fig. 34-83 shows one example of an artificial echo caused by poor loudspeaker layout.

One other source of artificial echoes, poorly understood by many designers, is related to feedback. Any sound from the loudspeakers that reaches the system microphone may be delayed by relatively long times in an outdoor system. This sound can be picked up by the microphone, reamplified, and emitted from the loudspeakers as an echo that cannot be distinguished from an echo created by a reflecting surface.

There are several ways to help solve this problem. One is to enclose the talker and microphone in a relatively soundproof room. In large outdoor stadiums, this is often the easiest way to avoid the regeneration type of echo. Another potential solution is to use a noise-canceling microphone located close to the talker's mouth. Close talking any system microphone will help, of course, because it allows reduction of system gain and an equal reduction of the reamplification of an echo. A noise gate set to turn off the microphone quickly after the talker stops talking may also help prevent regeneration echoes. This technique, however, will work well only in a situation where the microphone is relatively close to the talker's mouth.

The announcer on the field hears an echo that no one else in the stadium hears. That's because announcers first hear their own voice and then, delayed sometimes by as much as half a second or more, they hear their voice as an echo from the loudspeakers. This can be very confusing to an inexperienced announcer and may be the cause of the failure of many prepared speeches given to graduating classes in a football-field setting.

For the small football field with split bleachers, the distributed system concept may help since the sound is primarily aimed at the bleachers and not at the field. Even if there are some horns aimed at the field, they will normally be closer to the field than the horns on a typical scoreboard cluster. Thus, the signal delay and echo problem will be decreased.

Another partial solution to this problem is to give the announcer a local monitor loudspeaker (or headphones in an announce booth). The sound from this monitor will partially mask the echo from the cluster and may allow even an inexperienced talker to speak comfortably.

34.6.3.4 Dealing with Long Distances

Modular/concert-style line arrays may seem like an ideal way to throw sound over long distances outdoors. However, a typical line array has wide horizontal dispersion which may not be desirable. Also, few line arrays are rated for continuous outdoor exposure.

Thus, when a distributed system is not possible, a component cluster consisting of high-efficiency, high-power handling and narrow-coverage, constant-directivity horns is probably better than either line arrays or packaged loudspeaker systems for long-throw applications, Fig. 34-84.

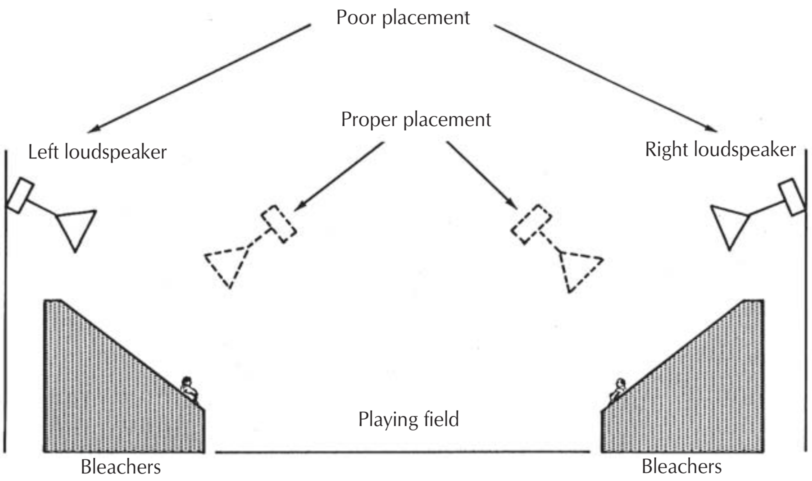


Figure 34-83. An artificial echo caused by poor layout of loudspeakers.

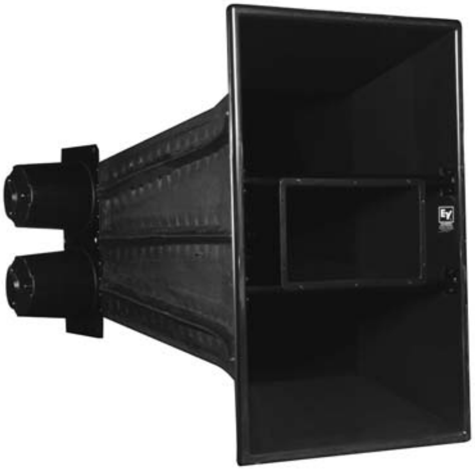


Figure 34-84. A large outdoor horn. Courtesy Bosch/Electro-Voice.

34.6.3.5 High Noise Outdoors (or Indoors)

How much L_P is required at the listener's position? The answer depends almost entirely on the expected ambient noise at the listener's location. Crowd noise in a sports stadium, for example, can easily exceed 90 to 95 dB L_P on the A scale. Noise from aircraft in an airport paging system or noise from race cars in a racetrack paging system may exceed this figure by a considerable margin. At a grand prix race, for example, spectators seated near the racetrack may be subjected to potentially ear-damaging short-term peak L_P levels of as high as 140 dB! In the case of the sports stadium, it may be possible to overcome the crowd noise (budget permitting). In the case of the racetrack, ask the announcer to avoid making announcements when the race cars pass the stands and to repeat announcements whenever possible. Also, put the announcer in an acoustically isolated booth so the announcement microphone does not pick up race noise as the cars pass the booth.

In the indoor system design equations in Section 34.2.3, a 15 dB SNR was assumed. For an outdoor classical music reinforcement system, this 15 dB SNR is still desirable. In many speech reinforcement systems, both indoors and out, a 15 dB SNR is simply not possible, and fortunately, it is not necessary. Except in the rare case where the frequency content of the noise is concentrated in the speech band, a 10 dB SNR will usually result in acceptably intelligible speech reinforcement. It is possible to achieve intelligible speech reinforcement with a SNR of lower than 10 dB but it is difficult to predict the success of such a system prior to its installation. When this type of system is contemplated, use a compressor to keep the voice level as constant as possible. Use a good-quality microphone with a bandpass filter to reduce unneeded low and high frequencies and avoid the use of a telephone handset as a paging microphone. Use equalization to boost the intelligibility frequencies slightly to further improve intelligibility. Use a smooth curve from about 1000 Hz to about 8000 Hz peaking in the 2000–4000 Hz bands.

Training the announcer in speech articulation will help, as will repeating announcements. In a plant or airport paging system with high noise levels, alerting the listener to an impending announcement by using a prepage tone can also help improve the effective intelligibility of a paged announcement.

34.6.3.6 Dealing with Varying Noise Levels

A manufacturing plant may have noise levels that vary widely with time and at different local areas within the building. Dealing with noise that varies in different areas is as simple as varying the quantity, type, and power delivered to the loudspeakers in the different areas. For noise levels that vary with time, there are two primary tools. If the noise varies predictably with time, as for an assembly line that is running or stopped on a predictable schedule, devices are available that vary the audio power fed to a loudspeaker line (usually by varying the input level to a power amplifier) depending on the time of day. Some of these devices will vary several different loudspeaker lines at several different times.

For noise levels that vary unpredictably with time, such as the crowd noise at a sports stadium or the noise of an airplane entering a waiting area at an airport, devices are available to measure the ambient noise level and adjust the paging level accordingly, Fig. 34-85. These devices work quite well, although they may have

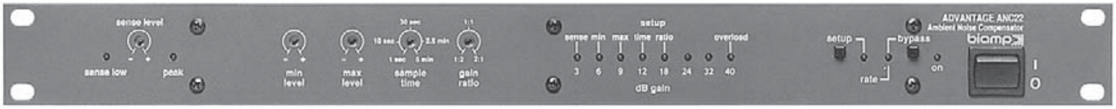


Figure 34-85. A device that increases or decreases sound system level with varying ambient noise. Courtesy Symetrix.

some trouble in a system that includes background music (the music may be interpreted as ambient noise or may prevent measurement of the noise), and they usually cannot adjust the paging level during a page.

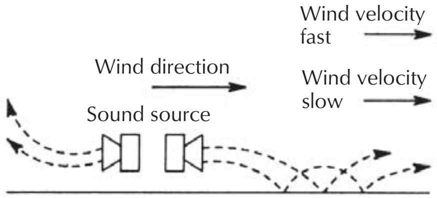
34.6.3.7 Thermal Layers of Air

Sound travels faster in warmer air. By itself, this fact has little significance for the sound system designer. Often, however, a layer of warm air will lie above or below a layer of cool air, and the difference in the speed of sound in these two layers can cause the sound to curve upward or downward toward the cool layer as shown in Fig. 34-86. This can be a help or a hindrance, depending on whether the cool layer is on the top or the bottom of the warm layer.

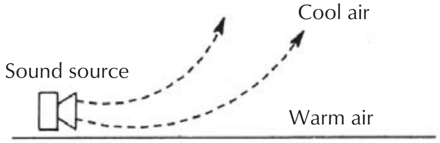
For example, in the early morning on a golf course, when the sun first comes up and begins to warm the air, the earth maintains a relatively cool layer near the ground. Thus, a sound wave will curve toward the ground, effectively hugging the ground, and can travel a great distance with seemingly little attenuation. Golfers at relatively great distances from each other can speak and be understood almost as if they were just a few feet apart. The same effect occurs over a quiet lake, even in the afternoon sun, since the lake will maintain a cool layer of air all day long. A wind, of course, will mix the layers of air and add noise so that the effect of thermal layers is lost.

In the early evening, when the sun begins to go down, the opposite situation occurs on a hot parking lot. The cool layer is now on top with a warm layer, maintained by the parking lot, on the bottom. The sound effectively curves upward toward the cool air and sound attenuation near the ground is effectively increased.

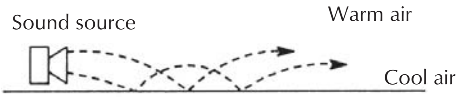
These phenomena can cause a paging system to work erratically, depending on the time of day. Large-area paging systems, over an airport runway, for example, are sometimes designed to overcome the changes caused by thermal layers. Elaborate systems can be designed to measure the temperature near the ground and above the ground and adjust the electrical input to each component of a vertical array of horns (or even mechanically re-aim the horns) to compensate for the effective curving upward or downward of the sound. In many smaller systems, however, the only necessary action is to test the system for satisfactory operation under worst-case conditions, which will probably be in the late afternoon on a hot day.



A. Wind mixes layers of air losing effect of thermal layers.



B. Warm air on the bottom causes sound to curve upward.



C. Cool air on the bottom causes sound to curve toward the ground.

Figure 34-86. Effect of thermal air layers on sound travel. Courtesy Bosch/Electro-Voice.

34.6.3.8 The Effect of Wind

Sound travels faster in the direction of the wind. Wind above the ground tends to be faster moving than wind near the ground. These factors combine to make the sound waves bend downward for a listener who is downwind from a sound source. Similarly, the sound waves will bend upward for a listener who is upwind from a sound source. Because this situation can vary unpredictably, wind can cause unpredictable changes in sound level and quality.

For outdoor concert systems with two stacks of loudspeakers (one on either side of the stage), a crosswind can cause very perceivable changes in phase cancellations at a given listener's position. In addition, because high-frequency horns are often very directional at higher frequencies, a small change in the direction of the sound can dramatically change the apparent high-frequency response to a listener.

While there are no permanent solutions to these problems, outdoor theater designers should be aware of them and should choose a site with as little wind as possible. Sound system designers may want to consider a distributed system approach to move the sound source closer to the listeners.

34.6.3.9 Dealing with Weather-Caused Deterioration

The outdoor environment is considerably more hazardous to the life span of a sound system than an artificially lighted, temperature-, and humidity-controlled indoor environment. Wind, rain, snow, ice, lightning, direct sun, widely varying temperature, salt air near the ocean, birds, squirrels, vandals, and pollution are just a few of the hazards faced by an outdoor sound system. Most of these hazards, of course, are felt by the loudspeaker system since the remainder of the sound system will usually be installed in a relatively safe indoor environment. However, avoid placing system electronics in a small room (equipment shack) that is exposed to direct sun and/or has poor ventilation. Also, protect or remove microphones, mixers, and so on that may be used in an exposed announcer's booth.

Use an effective lightning arrester above any system that might be exposed to lightning. Earth ground any metal gridwork and the metal chassis of loudspeakers to prevent static charge buildup, which can ultimately lead to arcing from the loudspeaker frame to the voice coil. Provide some type of static charge bleed to ground on any balanced, transformer-coupled loudspeaker line for the same reason. A pair of 1 MΩ resistors, one from each side of the line to ground, works well.

Choose a loudspeaker system intended for outdoor use. Outdoor loudspeaker enclosures and component horns should be coated with some type of weather-resistant finish such as epoxy paint, fiberglass, or the new specialty polyurethane coating now used by several loudspeaker manufacturers. Although a black or gray color is traditional, a white or other reflective color will help prevent heat buildup in hot sunlight. Fiberglass horns and enclosures should be painted white since the fiberglass resin will eventually evaporate in hot, direct sunlight if allowed to absorb heat.

Paper cone loudspeakers should be treated to resist damage from high humidity (this is a good idea in humid indoor environments, such as swimming pools, too). Simply spraying the cone with a waterproofing such as Scotchguard™ will help considerably and does not affect the performance significantly. The diaphragms of high-frequency drivers should either be made of a phenolic-type material or should be treated by the manufacturer to resist damage from humidity (most are). These treatments will also help prevent damage from pollution.

To protect against actual rain, install loudspeakers and horns pointing slightly downward if at all possible. If a horn must be pointed upward, use a curved adapter throat (available from the manufacturer) to point the driver downward. The adapter throat should have a weep hole in the bottom of the curve to allow water to drain out.

For additional protection against driving rain, some manufacturers use a layer of weather-resistant reticulated foam between two layers of grille cloth. Using a horn-loaded enclosure helps by recessing the loudspeaker cone. A layer of hardware cloth, or perforated aluminum, in addition to the grille cloth, can help prevent birds and squirrels from nesting in the enclosures and can help prevent damage from vandals throwing rocks, bottles, and so on. The hardware cloth can often be placed over the mouth of a horn, too.

Salt air can be extremely corrosive to metallic portions of loudspeakers, including metallic high-frequency driver diaphragms and metal horns. Use fiberglass or weather-resistant plastic horns, and coat low-frequency enclosures with epoxy paint or fiberglass. Consult with the manufacturer to choose low-frequency loudspeakers and high-frequency drivers that will resist damage from the salt air.

Especially when using a packaged loudspeaker system outdoors, consider the durability of the connectors. Conventional ¼ inch phone plugs, XLRs, or other indoor-type loudspeaker connectors will deteriorate outdoors from exposure to moisture and dust. Neutrik Speakon connectors are better but are not suitable for long-term outdoor exposure. The best solution is to eliminate the connector altogether and provide a direct cable connection into the loudspeaker system.

No matter how much care is taken in protecting outdoor loudspeaker systems, they will deteriorate faster than similar indoor systems. Thus, the system design should provide for easy access and repair, Fig. 34-87.

34.6.3.10 Simulating the Indoor Environment Outdoors

Many orchestras perform a yearly series of summer concerts in an outdoor amphitheater. Patrons enjoy picnics and music in an informal atmosphere. They may, however, notice that the orchestra's sound is not quite as rich or full as it would be during an indoor concert. The problem, of course, has nothing to do with the musicians and everything to do with the lack of beneficial reflections in the outdoor environment.

While the outdoor environment can probably never be as good for orchestral music as a great concert hall, steps can be taken to improve the richness of the sound.

Perhaps the most common approach is to reinforce the orchestra through a stage/shell-area loudspeaker system in conjunction with a series of remote loudspeakers (or small clusters) positioned to provide simu-

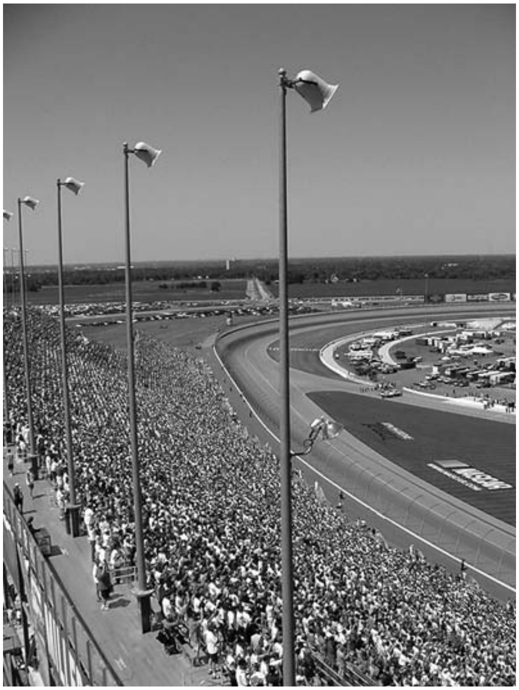


Figure 34-87. Outdoor-rated loudspeakers at a motor racetrack. Courtesy Community Professional.

lation of the reflections that would occur in a concert hall. Microphone usage and placement are similar to what would be used in the concert hall. An artificial reverberation device is used through the mixing console's effect send and return. The output of the mixer is fed to a multitap digital delay. The delay taps are selected carefully and fed to the individual loudspeaker systems to simulate the reflections that would be found in a concert hall.

Another approach is to use one or more semicircular rings of small loudspeakers at increasing radii into the audience. The rings may be divided into segments. Feed the rings from the output of the digital delay.

The process is not as straightforward as that of designing a simple sound reinforcement system and may take some experimenting. Nevertheless, it can significantly improve the sound of an outdoor orchestra and, thus, enhance the enjoyment of the audience.

34.6.4 Artificial Ambience

Large, multipurpose facilities designed for good speech intelligibility may not be desirable for musical performances because they do not have good early reflections and do not have a well-developed reverberant field. Some acoustical consultants now design such rooms with artificial ambience systems designed to add the early reflections and reverberation that are lacking in the natural acoustics.

Such systems consist of many microphones and loudspeakers and very sophisticated amplification, signal-processing, and control systems. These systems are complex and expensive but make it possible to hold a variety of events in a space and optimize the acoustics for each type of event. The services of a consultant, experienced in this type of system, are highly recommended.

34.6.5 Conference and Boardroom Systems

A conference room sound system has special problems because of the large number of open microphones, their proximity to loudspeakers, and the frequent added problems of ambient noise and uncooperative system users.

Some conference rooms are private; that is, they are designed exclusively for the use of the conferees. Other conference rooms include facilities for nonparticipating or infrequently participating observers, such as the stockholders, the press, or the public. Examples of the former are the conference rooms of privately held companies. Examples of the latter are the conference rooms of publicly held companies, unions, school boards, and governmental and other organizations. The sound systems installed in courtrooms are similar to those in conference rooms that include observers.

The primary difference between the two types of systems is in the additional loudspeakers (and roving microphones) needed for the second type of system. Both systems, of course, often include audio recording and playback capabilities and may include video playback and recording, audio or audio and video conferencing capabilities, computer graphics display capabilities, and so on.

34.6.5.1 Local Sound Reinforcement in a Conference Room

Microphone choice is important to the success of a conference room system. One school of thought favors boundry microphones placed on the table in front of each participant. A boundry mic is inconspicuous and does a good job of picking up the direct sound from the talker's voice. Because the boundry mic element is less than a millimeter from the table surface, it does not pick up reflections from the table in the frequency range of interest. Thus, technically, this is a good choice.

The problem with boundry mics is the users. If the mics are loose, the users will push them out of the way or turn them around. If they are fixed in position, users will lay paperwork or books on top of them, spill water or coffee on them, and otherwise abuse the microphones. Boundry mics also do a very good job of picking up the noise of paper shuffling on the table.

Thus, another school of thought suggests installing gooseneck mics at each talker's position. The gooseneck mic puts the mic element up near the talker's mouth so it does a good job of picking up up the direct sound. Unfortunately, the first reflection from the table also enters the gooseneck mic causing undesirable comb filtering. Also, it's easy for the user to push the gooseneck mic out of the way.

Fortunately, gooseneck mics seem to attract users who often pull them close and talk directly into the microphone. When this happens, the direct sound is much louder than the first reflection from the table so the reflection is of little consequence. Thus, the gooseneck mic may work better than the boundry mic even though it is technically inferior (because of the reflection from the table).

In some conference rooms, the designer may need to try both microphone types to see how the users react. It may also be a good idea to give each user a lighted mute button for private conversations with nearby conferees.

Conference room loudspeaker systems are normally laid out in a distributed fashion. To help avoid feedback from multiple open microphones and to help avoid pickup of paper shuffling and other ambient noises, an automatic mixer is commonly used (see Section 34.4.3). Often, the chair position is given an automatic priority switch so that when the person at the chair position speaks, all other microphones are turned off (this may also be accomplished by a manual priority switch).

Another way to help avoid feedback is to use the logic outputs on the back of many automatic mixers to switch off the loudspeaker directly above the head of the person talking. This approach, however, can hinder the kind of back-and-forth conversation where more than one person talks at the same time and there are frequent interruptions. It should be noted that, in a conference room, this kind of conversation is difficult with any system.

A better way to reduce feedback is to use a mix-minus system as discussed in Section 34.4.3.3. To make this work for a large conference table, design the system so that each loudspeaker receives its own customized mix including all microphones except the one nearest the loudspeaker. Taper the mix for microphones located farther away from the local microphone as shown in Fig. 34-41. Loudspeakers for inactive participants can be on all the time unless a roving microphone is being passed around. In this case, the level to the local loudspeakers may need to be reduced somewhat, and, of course, the feed from the roving microphone to the conference table should be fed to all loudspeakers at equal levels.

34.6.5.2 Recording a Conference

The multiple-microphone setup of a conference can be a problem for recording. All microphones can simply be mixed and fed to a single channel of an audio recorder, but the multiple microphones may add unwanted ambient noise. Using the output of the automatic mixer helps avoid this problem. Use one of the matrix outputs when designing a mix-minus system. A courtroom-style, multichannel, logging recorder may also be used to record a number of individual microphones on separate channels. This is especially useful when a transcript of the conference must be made later. In some systems, the chair is given a switch to pause the recording for off-the-record conversations.

34.6.5.3 Audio and Video Teleconferencing

Although speaker-phone conferences over normal telephone lines are still commonplace, most video/audio teleconferencing now takes place over the Internet or dedicated wide area network (WAN). Whatever method is used, the characteristics of the transmission path must be considered. In particular, it is possible for feedback and echoes to occur over a complex path including the transmission line, the local microphone(s) and loudspeakers and the remote microphone(s) and loudspeakers.

A digital echo canceller is built in to every modern video/audio teleconferencing system. These devices help solve the problems of echoes and feedback in a teleconference. To optimize the echo canceller, however, the local audio systems in both rooms must be properly designed.

In most cases, this design is very similar to the design of a conference room designed for local sound reinforcement. Choose microphones carefully as described in Section 34.6.5.1. Lay out a distributed loudspeaker system. Use an automatic mixer with matrix output to create a mix-minus system. Do this in both conference rooms.

In operation, the echo canceller takes some time to adjust itself to the system. After it optimizes its operation, the system should be relatively free of echoes or feedback. The echo canceller must reoptimize itself when there is any change in the system so try not to move any microphones once a conference has started.

34.6.6 Sound Masking (Noise Masking), Speech Privacy Systems

Sound masking is an electronic system that creates a low-level "masking sound" to improve speech privacy and mask irritating noises in a work space. A well-designed sound masking system provides protection for confidential conversations and creates a more pleasant work environment. A sound masking system may also be called a noise masking or white noise system.

Open-plan offices are the most common application for sound masking. Doctors' exam rooms, and other facilities that fall under HIPPA privacy regulations, are also good places for sound masking. Sound masking may also be used in courtrooms, law offices and in high-security government or industrial facilities.

34.6.6.1 Criteria for the Environment

A limited range of environments is suitable for sound-masking. The environment must be relatively quiet since the masking sound will need to be louder than the noise to be masked. A general criterion for ambient noise level is about NC-35 (noise criteria curve number 35) or about 45 dB on the A scale. The environment must have a low-reverberation time and as few hard, reflecting surfaces as possible. A well-designed open-plan office will usually meet these criteria.

34.6.6.2 Mechanical Noise Reduction

Noise from computer printer areas, copying machines, and so on should be reduced by mechanical and acoustical means since the sound masking system cannot be expected to overcome these relatively high-level sources. In addition, mechanical barriers in the form of open-plan office-type acoustical dividers are normally installed between offices to help attenuate noises and speech from point to point. Nonreflective ceilings (preferably dropped acoustical tile) and carpeted floors are almost mandatory. Without these mechanical aids, the sound masking system cannot perform its function successfully.

34.6.6.3 Masking Sound

Once higher-level noises have been reduced and an acceptable acoustical environment has been created, low-level noises become more irritating, and speech privacy becomes more important. Masking sound, created by a sound masking, speech-privacy system, can help solve these problems, Fig. 34-88.

Masking sound systems should be located in the vicinity of the listeners, not near offending noise sources. Masking systems must be completely unnoticeable to the listeners. That is, no one should know there is a sound-masking system in operation unless someone turns it off. The reason for this, of course, is that the purpose of the system is to help reduce distracting noises and to aid speech privacy. If the system itself becomes a distraction, one of its primary purposes has been defeated.

The masking sound itself is created by an electronic random-noise generator similar to that used for equalization, and it is fed through an equalizer to a set of power amplifiers and loudspeakers. The loudspeakers are normally hidden in the ceiling plenum, above the acoustic tiles.

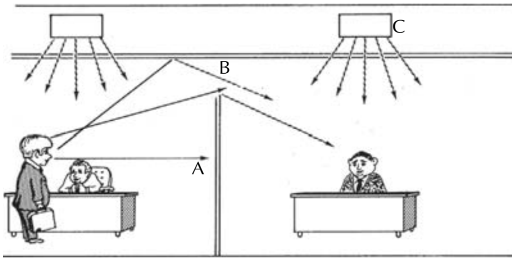


Figure 34-88. Factors affecting speech privacy. Courtesy Atlas Sound.

34.6.6.4 Criteria for the Loudspeaker System

To meet the primary goals of sound masking and speech privacy and to remain unnoticed by listeners, the loudspeaker system must produce random noise that does not change in level as a listener moves from place to place within the environment. A traditional downward-facing distributed system could achieve this goal, but only with an extremely high density of loudspeakers. Thus, an upward or sideways facing system is more common, Fig. 34-89. The upward and sideways systems use mechanical structure in the ceiling plenum to reflect and help randomize the noise distribution. In a typical open-plan office with a ceiling height of 8–9 ft, the loudspeakers would be spaced in a square or hexagonal pattern with approximately 10–20 ft spacing between individual loudspeakers. Care must be taken to avoid placing a loudspeaker too near a hard reflecting object, such as an airduct, which might cause an audible hot spot in the room below. Various ceiling materials and baffles are available to help diffuse the masking sound, and several manufacturers produce special loudspeaker-enclosure combinations designed specifically for sound-masking to allow the upward or sideways facing orientation.

34.6.6.5 Criteria for the Amplification System

A standard pink-noise generator, of the type used for sound system equalization, is acceptable provided its output is random enough. Many equalization noise generators utilize digital circuitry, with only pseudorandom noise. Listeners may become aware of this type of generator making it unacceptable. Noise generators specifically designed for sound masking systems are available and are worth any extra cost. The noise generator, like the power amplifier, must be highly reliable. Redundancy can be achieved by mixing two noise generators through a passive mixing network. If one generator fails, the overall noise level will drop about 3 dB, but the system will continue to operate.

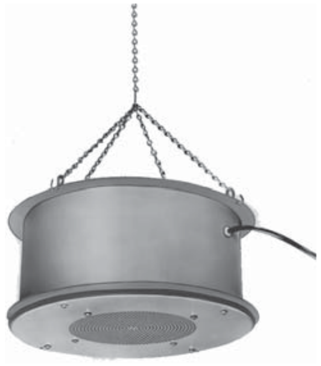
34.6.6.6 Criteria for the Equalizer

Because the shape of the final frequency-response curve is critical, and standard masking system curves are specified in ⅓-octave intervals, a ⅓-octave equalizer should be employed. If two, redundant, active equalizers are used, the filter and gain settings on both must be exactly equal.

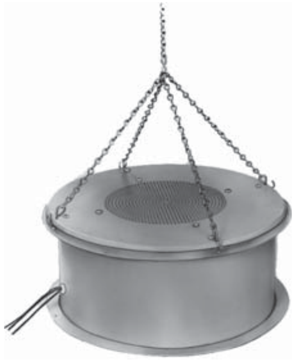
34.6.6.7 A Second System

A more random (and thus more effective) system design can be achieved by utilizing two separate noise generators feeding two equalizers and two power amplifiers (or banks of power amplifiers in a large system) as shown in Fig. 34-90. Rather than feeding zones of loudspeakers in separate areas, these amplifiers feed adjacent loudspeakers in the same zone in a checkerboard pattern. This plan also produces a higher level of redundancy since failure of one amplifier or noise generator will produce only a 3 dB drop in overall masking level. Thus, this system can be installed with no backup amplifiers if desired. Alternately, a single backup amplifier can be installed in the system rack, ready to replace any single amplifier failure.

A. Sideways.



B. Downwards.



C. Upwards.

Figure 34-89. Sound masking loudspeaker shown in three hanging orientations. Courtesy Atlas Sound.

34.6.6.8 Masking Plus Paging or Background Music

Background music or paging may be added to a masking system using the same amplifiers and loudspeakers. Intelligibility of the paging may suffer because of the placement of the loudspeakers. (High frequencies will be attenuated.) A separate equalizer should be used for

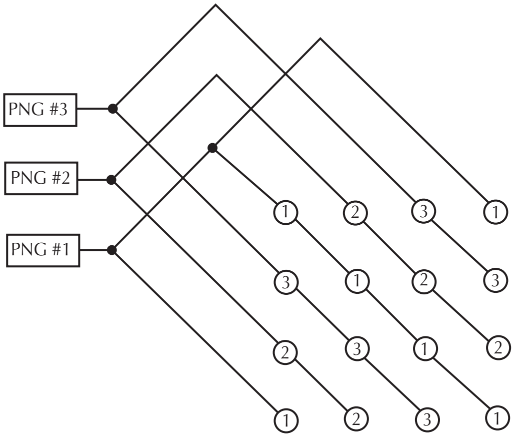


Figure 34-90. Checkerboard loudspeaker pattern with dual masking generators. Courtesy Atlas Sound.

the paging; if all three functions are to be included, another equalizer should be used for the background music. A multichannel DSP device may be used if the reliability is sufficient. The system power amplifier (and loudspeakers) must be capable of handling simultaneous inputs from all sources with adequate head room. The masking noise must not be turned off or attenuated during a page, since this will cause listeners to become aware of the system, detracting from its effectiveness. It is generally better to separate the masking system from the paging and music system, and design the latter as a standard distributed system.

34.6.6.9 Adjusting the Installed System

Both the level and frequency response of the masking system must be properly adjusted. Perform the adjustments when office workers are not present. Adjust the masking noise level for the degree of speech privacy required keeping the masking sound as low as possible consistent with speech privacy requirements. Ideal levels are between 45 dBA and 48 dBA. When needed, office workers will often tolerate masking-sound levels of as high as 52 dBA, but higher levels will defeat the purpose of the masking system by making it into an irritation itself. If acceptable speech privacy is not achieved at this masking sound level, alternate mechanical and acoustical means should be employed.

The frequency-response curve must also be carefully adjusted to conform to the window curve shown in Fig. 34-91. This curve includes the effects of existing mechanical noise sources such as air-handling systems, and these sources normally contribute the bulk of the noise energy below about 250 Hz. Both frequency response and sound pressure level must be measured at multiple points in the room and variations of more than about ±2 dB can mean degraded system effectiveness.

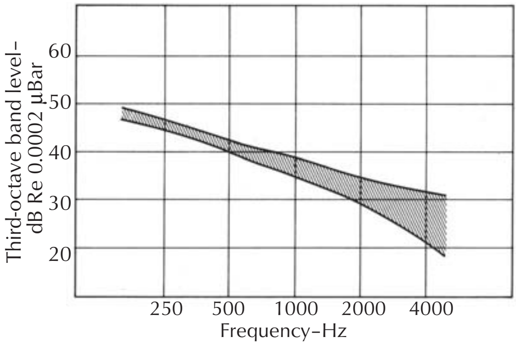


Figure 34-91. Range of typical masking sound spectra. Courtesy Atlas Sound.

If other areas of the building, especially on the same floor, will not have a sound masking system, plan the system so that a transition zone can be achieved where the masking sound gradually dies out as a listener moves from one area to another. Adjust this transition zone so that a listener walking from one zone to the other notices only a subjective natural change in sound as might be expected in walking from one area to another.

34.6.6.10 Objective and Subjective Methods for Measuring System Effectiveness

Effective speech privacy is the most important goal of a sound masking/speech-privacy system. Speech privacy is most often specified and evaluated in terms of the Articulation Index (AI) as specified by ANSI Standard S3.5. The Articulation Index measures speech intelligibility as a weighted sum of signal-to-noise ratios in multiple frequency bands. An AI of 1.00 is considered to be excellent intelligibility or no privacy. An AI of 0.00 is considered to be bad intelligibility or confidential privacy. An AI of 0.05 to 0.19 indicates normal privacy. An AI of 0.20 to 0.32 indicates marginal privacy. AI can be evaluated by measuring signal to noise (speech level to background noise level) in the indicated bands, multiplying by the given weighting factors and adding the results. This is normally performed by an acoustical consultant.

A subjective evaluation of system effectiveness should be performed by a jury of at least three listeners. Position each listener (one at a time, independently) on the other side of an acoustic barrier from a talker (this simulates the normal usage of the office area). Measure the talker's voices and have them talk louder or quieter until the talker level is about 60–65 dBA. Adjust the masking sound level until the listener agrees that the speech from the talker is audible but not intelligible. Repeat the process for the other two listeners. The final system level should be the highest of the three levels. This minimal privacy level should not be exceeded if at all possible.

34.6.6.11 Conclusion

Sound masking, speech-privacy systems are related to other sound systems only in the fact that they use many of the same components. Their purposes and design are obviously very different from each other. Masking systems are commonly specified for public buildings. Such specifications are drawn up in great detail and requirements for installation and performance documentation are strict. When a design must be performed from scratch, the help of an experienced consultant will be valuable.

34.7 Computers and Software for Sound Reinforcement Systems

Computers and software programs impact the audio world in several distinct areas as briefly discussed here. For additional details, see Chapters 31—DSP Technology, 14—Transmission Techniques: Wire and Cable, 15—Transmission Techniques:Fiber Optics, 25— Consoles, 29—MIDI, 38—Virtual Systems, 9—Modeling and Auralization, and 46—Audio Measurements.

34.7.1 Digital Audio Devices

Eventually, most audio products will be based on digital technology. DSP signal-processing systems, as discussed in Section 34.4.5, have already become mainstream products that are extremely useful to the sound system designer. Several manufacturers are now offering fully digital audio mixing consoles. Even analog mixing consoles enjoy important digital features including mute groups and VCA groups.

34.7.2 Audio Signal Transmission via Computer Networks

Cobranet™, developed by Peak Audio, allows transmission of multiple audio signals over a conventional ethernet computer network. Other proprietary digital audio transmission systems are available. Eventually, standards of this type will enable audio signals to remain in digital form from the microphone element to the amplifier input and possibly even to the loudspeaker voice coil.

34.7.3 System Control and Monitoring by Computer

Safety regulations may require that paging systems used for life safety purposes, such as emergency evacuation, be supervised. This usually means a computerized means of monitoring the impedance of the loudspeaker line. If the impedance changes more than a specified amount the computer assumes there is a loudspeaker failure or line short and it signals trouble on the line.

Non-life-safety sound systems may not require computer control and, in many cases, would not benefit from it. Larger, more complex systems, however, can benefit greatly from a central computer set up to monitor and control the system.

Consider a large airport paging system. Paging messages originate in individual gate areas, baggage claim areas, a central paging room, one or more security areas, and even from outside the buildings. Listeners are located in many different areas as well.

Paging messages range from "Mr. Smith, meet your party at Baggage Claim Area 2." to gate change announcements, flight change announcements, and emergency announcements.

Each message must be routed to the appropriate areas in the appropriate terminals. A priority system must be established for two or more messages that originate at approximately the same time. Emergency priority must be given when needed. In addition, the system equipment may be widely dispersed and must be monitored for possible failures. A computerized system, like that shown in Fig. 34-92, can perform all of these functions and more.

34.7.4 Computer Aided System Design

As discussed in Section 34.3.2.9.2, software programs like EASE and Modeler are extremely valuable system design aids. Less complex software, such as the spreadsheets offered by Syn-Aud-Con can also aid the system designer. Other software, such as Stardraw, helps the designer perform system layout and documentation.

34.7.5 Computer Aided Measurement Systems

As discussed in Chapter 50, computer aided measurement systems like SMAART from SIA Software; TEF from Gold Line; EASERA, an acoustic measurement program offered by the developers of EASE; MLSSA from DRA Laboratories; SIM from Meyer Sound; and others have greatly advanced the state of the art in acoustic and electronic measurements.

Figure 34-92. A computerized airport paging system. Courtesy IED Audio.

References

1. H. F. Hopkins and N. R. Stryker, "A Proposed Loudness-Efficiency Rating for Loudspeakers and the Determination of System Power Requirements for Enclosures," *Journal of the Proceedings of the IRE*, March 1948.
2. V. M. A. Peutz, "Articulation Loss of Consonants as a Criterion for Speech Transmission in a Room," *J. Audio Eng. Soc.*, Vol. 19, No. 11, December 1971.
3. W. Klein, "Articulation Loss of Consonants as a Basis for the Design and Judgment of Sound Reinforcement Systems," *J. Audio Eng. Soc.*, Vol. 19, No. 11, December 1971.
4. Don Davis and Eugene Patronis, *Sound System Engineering, Third Edition*, Focal Press, 2006.
5. E. Lethert, "Calculating N for Mixed Acoustic Sources," *Synergetic-Audio-Concepts Tech Topics*, Vol. 5, No. 5, January 1978.
6. E. T. Patronis and C. Donders, "Central Cluster Design Technique for Large Multipurpose Auditoria," *J. Audio Eng. Soc.*, Vol. 30, No. 6, June 1982.
7. Don Keele and Don Davis, "More on Centers and Origins; Alignment Guidelines," *Syn-Aud-Con Tech Topics,* Vol. 16, No. 7, Summer 1989.
8. John Eargle and Chris Foreman, *Audio Engineering for Sound Reinforcement*, Hal Leonard, 2002.

Organizations and Web Sites

There are a multitude of Web sites dedicated to professional audio. Some are hosted by professional audio manufacturers. Others are hosted by professional audio trade magazines. Some are hosted by industry organizations. A few are independent. Here are a few that are recommended by this author. Many of these Web sites offer links to other good Web sites so the possibilities for good pro audio reference material on the Web are almost unlimited.

Synergetic Audio Concepts, www.synaudcon.com. Founded in the early 1970s by Don and Carolyn Davis, Synergetic Audio Concepts is the best known independent organization dedicated to audio education. Current owners, Pat and Brenda Brown offer regular technical seminars and maintain an excellent Web site.

Trade Organization Web Sites

Audio Engineering Society, www.aes.org

Acoustical Society of America, www.acoustics.org

Infocomm, www.infocomm.org

NSCA, www.nsca.org

NAMM, www.namm.org

Rane Professional Audio, www.rane.com. Rane offers a well-respected series of technical papers called “Rane Notes” and a book called *Pro Audio Reference*.

Jensen Transformers, www.jensen-transformers.com. Bill Whitlock of Jensen Transformers is the author of Chapter 32, “Grounding and Interfacing” of this *Handbook for Sound Engineers*. His company maintains an excellent web site with numerous white papers and references.

JBL Professional, www.jblpro.com. JBL offers a good selection of technical documents including the historic *JBL Professional Sound System Design Manual* and a group of AES papers.

Pro Sound Web, www.prosoundweb.com. Excellent independent site with technical articles, forums, and an unequaled set of pro audio Web links originated by Ken Berger of EAW (links.prosoundweb.com).

Bosch/Electro-Voice, www.electrovoice.com. The original *PA Bible* is still available on this Web site and Electro-Voice has recently added to this well-known and excellent reference source.

Pro AV Magazine, www.proavmagazine.com. This is one of the best trade magazine Web sites for technical information. Click on “Resources.”

Bibliography

Journal of the Audio Engineering Society. Available to members only. AES Membership Information, 60 East 42nd Street, New York, NY 10017, or at www.aes.org.

D. L. Klepper, Ed., “Sound Reinforcement, An Anthology,” New York: *Audio Engineering Society.*

The PA Bible, A series available from Bosch/Electro-Voice.

J. R. Prohs and D. E. Harris, “An Accurate and Easily Implemented Method of Modelling Loudspeaker Array Coverage,” Audio Engineering Society preprint 1941(A-8), October 1982.

H. G. Smith, “Acoustic Design Considerations for Speech Intelligibility, *J. Audio Eng. Soc.*, Vol. 29, No. 6, June 1981.

Sound System Design Reference Manual and other publications available from JBL Professional, 8500 Balboa Blvd., P. O. Box 2200, Northridge, CA 91329.

Syn-Aud-Con Tech Topics, a continuing series of technical newsletters published by Syn-Aud-Con at www.syn-aud-con.com.

Technical Letters. A continuing series of technical and application notes originally published by Altec Lansing. Many are still available at http://alteclansing unofficial.nlnet.net/publications/techletters/index.html.

T. Uzzle, "Loudspeaker Coverage by Architectural Mapping," Audio Engineering Society preprint 1786(G-3), May 1981.

Acknowledgments

The author extends his personal thanks to the following individuals for their help with this chapter of the *Handbook for Sound Engineers*:

Don and Carolyn Davis, Founders, Syn-Aud-Con.

Glen Ballou, Editor, *Handbook for Sound Engineers*, Guilford, CT.

Chapter 35

Computer Aided Sound System Design

by Dr. Wolfgang Ahnert, Stefan Feistel and Hans-Peter Tennhardt

---

Introduction 1339
35.1 Sound Design Basics for Acoustic Simulation 1339
  35.1.1 Measurement and Planning Methodology with Physical Models of Large Auditoriums 1340
    35.1.1.1 Fundamentals 1340
    35.1.1.2 Balance Investigations of Music Performances 1341
  35.1.2 Building a Computer Model, Entering Room Data 1342
  35.1.3 Acoustic Sources and Loudspeaker Systems 1343
    35.1.3.1 Natural Sources 1343
    35.1.3.2 Loudspeaker Types 1344
35.1.3.2.1 Point Sources 1344
    35.1.3.3 Loudspeaker Line, Sound Column and Line Arrays with In-Line Arrangement of Radiators 1350
  35.1.4 Wall Materials 1354
    35.1.4.1 Absorber Data 1354
    35.1.4.2 Scattering Data 1354
    35.1.4.3 Diffraction, Low-Frequency Absorption 1354
  35.1.5 Receivers and Microphone Systems 1355
    35.1.5.1 Human Ears 1355
    35.1.5.2 Microphones 1355
35.2 Transducer Data for Acoustic Simulation 1356
  35.2.1 Simulation of Loudspeakers 1356
    35.2.1.1 Simulation of Point Sources 1357
    35.2.1.2 Simulation of Complex Loudspeakers 1360
35.2.1.2.1 Modeling by Means of a DLL Program Module 1360
35.2.1.2.2 Modeling by Means of a GLL Data File 1362
    35.2.1.3 Background of Simulation and Measurement 1362
35.2.1.3.1 Resolving the Far-Field Problem 1363
35.2.1.3.2 Acquisition and Interpolation of Complex Data 1363
35.2.1.3.3 Configurable Loudspeakers 1365
35.2.1.3.4 Shadowing and Ray Tracing 1367
35.2.1.3.5 Additional Notes 1367
  35.2.2 Receiver Simulation 1368
    35.2.2.1 Simulation of the Human Head 1369
    35.2.2.2 Simulation of Microphones 1369
35.3 Tools of Simulation 1370

- 35.3.1 Room Acoustic Simulation . . . 1370
  - 35.3.1.1 Statistical Approach . . . 1370
  - 35.3.1.2 Objective Room-Acoustic Measures . . . 1371
- 35.3.2 Ray Tracing or Image Modeling Approach . . . 1371
  - 35.3.2.1 Preliminary Remark . . . 1371
  - 35.3.2.2 Image Modeling . . . 1371
  - 35.3.2.3 Ray Tracing . . . 1372
  - 35.3.2.4 Cone Tracing . . . 1372
  - 35.3.2.5 Pyramid Tracing . . . 1372
  - 35.3.2.6 Room-Acoustic Analysis Module-(AURA) . . . 1373
  - 35.3.2.7 Features of All These Methods . . . 1374
- 35.3.3 Auralization. . . . 1374
- 35.3.4 Sound Design . . . 1374
  - 35.3.4.1 Aiming . . . 1374
  - 35.3.4.2 SPL Calculations . . . 1376
  - 35.3.4.3 Time Arrivals, Alignment . . . 1378
  - 35.3.4.4 Mapping, Single-Point Investigations . . . 1378

35.4 Verification of the Simulation Results . . . 1379

References . . . 1382

Introduction

For more than 2000 years acoustic phenomena have been perceived and manipulated subjectively. Reference can be made in this context to Marcus Vitruvius, the ancient Roman architect who described at this early time the application of acoustic laws in theatrical spaces. But only since the end of the medieval times and particularly during the last century acoustics has developed into an independent science.

Highlights on the way to a scientifically calculated design:

- Roman/Greek times, medieval times: knowledge based on experience and first trial and error reports—e.g., by the Roman architect Vitruvius, 15 BC.
- Since the end of 18th century: Theoretical investigations—e.g., Chladni, 1810, or in 1875, Lord Rayleigh, Professor Helmholtz.
- Since 1900: room acoustical basics, Professor Sabine/United States 1923; radiation of sound, H. Stenzel/ Germany 1930, and H. F. Olson/United States 1947.
- By 1935: measurement in models and "auralization" in physical models, Professor Spandöck München, Professor Reichardt, Dresden/Germany.
- Since 1965: computer-model investigations, Professor Krokstad, in Trondheim/Norway, afterwards many similar works have been done.
- Since 1995: auralization by means of computer models has been introduced.

A measured sound-field structure in 3D, so-called waterfall form (decay of sound energy as a function of time and frequency) is shown in Fig. 35-1. The sound level is marked on the ordinate, the frequency (range of frequency 63 Hz–8 kHz) on the abscissa, and the time on the third axis (0 ms–direct sound to 3.5 s reverberation).

These sound-field structures are depending on listener locations. In the old days a wanted sound decay for concerts or for speech transmission was generated by changing the primary or secondary structure of a room, see Chapter 7.3.2. Now with sound systems it is possible to generate any sound fields subjectively desired.

Today we can derive the basic items in sound design. Listening comfort and intelligibility are influenced by:

- Reverberation time and volume.
- Early and late reflections.
- Ambient noise level.
- Directivity of the loudspeakers or new array constructions.
- New loudspeaker types.
- Interference effects.

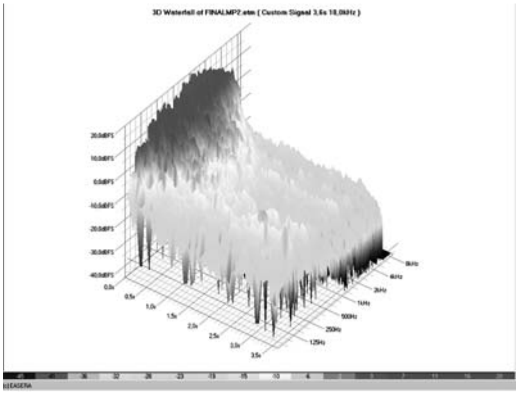


Figure 35-1. Typical waterfall presentation.

Some basic measures of the performance of a sound system are:

- Intelligibility *Alcons*, C_t, RASTI, etc.
- Loudness in dB ($\boldsymbol{SPL}_{tot}$).
- Direct sound in dB (SPL_{dir}).
- Frequency response in ±dB from flat.
- Coverage in ±dB from even.

The goal of modern sound design is to calculate in advance the complete sound field structure in a hall or in open spaces by means of Computer-Aided Acoustic Design (CAAD) programs enabling you to prevent any surprises that become evident after a sound system has been installed. You just describe in advance the expected properties of your sound system and the overall acoustic properties of your room using the new sound system.

The following considerations include personal experiences of the authors, especially with the CAAD program EASE,[1] but features of other programs are also explained.

35.1 Sound Design Basics for Acoustic Simulation

Application of physical or computer models for acoustic and sound system design:

Prior to 1965. Physical models since the 1930s and after WWII mainly in selected cases in research centers by means of huge computers.

Since 1970. The programmable pocket computer offered the first algorithms for acoustics.
1980–1985

Since 1981. The PC and PC-XT have been available.

1984: Reverberation time and intelligibility calculations for simple rooms:
CADP / JBL.
TEF-10-Analyzer. First coverage plots became reality.

1985: PHD Program. By Prohs/Harris in spring: first version for TEF Analyzer:

1. Room-acoustic calculations like different reverberation times.
2. Loudspeaker cluster design.
3. Power calculation for horn radiators and corresponding drivers.
4. *Alcons* calculations by Peutz.

1986: BOSE-Modeler. First full-graphic CAD Macintosh-based program, Version 1, 1986, by K. Jacob, T. Birkle/Bose/United States.

1987: Acousta-CADD. First full-graphic CAD MS-DOS-based program, Version 1, by A. Muchimaru, Altec Lansing/United States.

1990: EASE. Full-graphic CAD MS-DOS-based program with pop-up menus, Version 1, 1990, by ADA, Germany.

1991: CADP2. Full-graphic CAD Windows 3.1-based program, by JBL/United States.

1996: ULYSSES. By IFB/Germany (P. Hallstein).

1997: CADP2. Further development stopped.

1999/2001: EASE for Windows. By Ahnert Feistel Media Group.

Room Acoustics Programs

1988: *CATT-Acoustic*. Dalenbäck/Sweden, Version 1, now Version 8.0.

1991: *ODEON*. Naylor & Rindel/Denmark, Version 1, now Version 9.0.

1994: *RAMSETE*. Farina/Italy, Version 1, ***n***ow Version 2.5.

1998: *CAESAR*. Vorländer/Schmitz/Aachen/Germany, Version 0.12, 2001 Version 0.20.

2002: *AURA*. (Room acoustic module) in **EASE** Version 4.2, now Version 2.2.

The programs printed in italics are subject to constant advancement.

35.1.1 Measurement and Planning Methodology with Physical Models of Large Auditoriums

by Hans-Peter Tennhardt

35.1.1.1 Fundamentals

The room impulse response is obtained in a reduced model of the auditorium interior by applying the corresponding scale-model laws based on the constant ratio between the geometrical dimensions L of the room and of the sound wavelength λ in the model scale (index M) and in natural scale (index N):

$$\begin{aligned}\frac{L}{\lambda} &= Const\\ &= \frac{L_N \times f_N}{c_N}\\ &= \frac{L_M \times f_M}{c_M}\end{aligned} \qquad (35\text{-}1)$$

where,
c is the speed of sound,
f is the frequency.

If the scale-model test is carried out in the same sound propagation medium, then $c_N = c_M$ and so Eq. 35-1 becomes

$$\begin{aligned}p &= \frac{L_N}{L_M}\\ &= \frac{f_M}{f_N}\end{aligned} \qquad (35\text{-}2)$$

i.e., the measurements are carried out in a frequency range that exceeds the original frequency range by the factor p (reduction scale 1:p).

A favorable compromise regarding model size and accuracy of reproduction is given with a reduction scale of 1:20, but scales between 1:8 and 1:50 are feasible depending on model size or frequency range to be studied. The sound impulse is irradiated from the location of the sound source (e.g., stage, orchestra pit, loudspeaker). The acoustical response of the room to the emitted signal is simultaneously registered at receiving positions (audience area, platform, stage) by special electroacoustic transducers (microphone, dummy head with ear simulator). The transfer function between transmitting and receiving locations is calculated from the obtained room impulse response. Very often a spark discharge generator is used as a sound transmitter in air (nowadays electronic MLS scale radiators are also used). With an impulse width in the model of 80 µs, it is possible to resolve path differences equaling 60 cm in the original room. The reproducibility of the maximum sound pressure is below ±0.2 dB.

Special model sound sources enable simulation of a talker or a singer, of an orchestra as nondirectional sound source in the center of the same, of orchestral instrumental groups (see Section 35.1.2.2), and of loudspeaker lines with variable directivity characteristics.

Preferably the registration of the room impulse response is dual-channel by the microphones of the dummy head at listener seats, which are representative for determined seating groups so that the binaural head-related listening parameters of the human auditory organ are optimally reproduced. In a model of scale 1:20 the diameter of this miniature dummy head must be about 11 mm, Fig. 35-2.

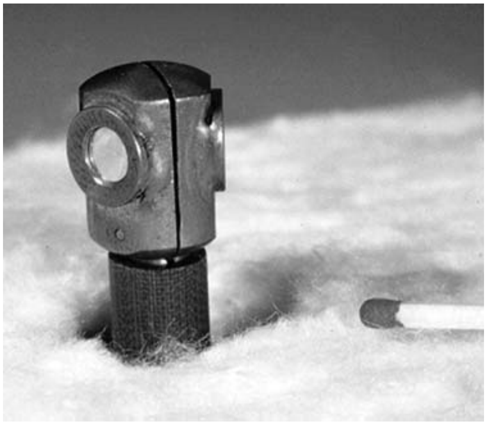


Figure 35-2. Model-size dummy head for measurements at a scale of 1:20.

The investigated frequency range lies between 5 kHz and 200 kHz in the scaled model, which corresponds to 250 Hz to 10 kHz in the original. Structures whose linear dimensions fall below approximately 8 cm in the original room are not reproduced in the model. Also sound absorbers and wall impedances below the studied frequency range are not considered in the model tests.

For better access to the models during the measurements, these should be carried out in air under normal pressure. Owing to the excessive atmospheric absorption occurring at the model frequencies there occurs a faulty momentary value of the sound pressure that is mathematically corrected by a real-time compensation. Without these mathematical corrections the measurements must be carried out in a nitrogen environment instead of air, in which case the drawback consists of bad access to the model.

All physical acoustic phenomena such as diffraction and scattering are represented in a frequency-true fashion.

The obtainable accuracy is presently still superior to that of a computer simulation. The method is capable of providing answers to questions concerning balance investigations in rooms for music performances, see Section 35.1.2.2, the influence of electroacoustical components on room-acoustical parameters, and the directional effect of wall and ceiling structures prepared as scaled models.

By using original sound source simulations (talker, singer, nondirectional sound source, orchestra instrumental groups, loudspeakers) and a dummy head as a receiver unit, the described measuring procedure is applicable also for original rooms.

35.1.1.2 Balance Investigations of Music Performances

The scaled-model simulation of an orchestra can in first approximation be realized by a nondirectional sound irradiation from the center of the same. A more detailed simulation is necessary, however, if one wants to have information about the influence of a room on the balance of the different instruments at the listener's seat. A useful approximation can be obtained by a simulation of orchestra instrumental groups in which their sound spectra are based on the frequency response chiefly reproduced in music presentations. Their directional characteristics are derived from the usual playing posture.[2]

The simulated orchestra is subdivided into four instrumental groups, in which the percussion instruments, in view of their considerable loudness and adaptable style of playing, may be left out of consideration:

- String instruments (St)
- Woodwind instruments (Wo)
- Brass instruments Bl (Bi)
- Bass instruments Ba (Ba)

To this may be added the electroacoustical model transducer of a singer/talker (S).

The scaled-model simulation comprises the impulse excitation by a spark-gap generator provided with a shading reflector of defined sound attenuation so as to align it with the directional characteristic of the instrumental group in question.[3,4] The electroacoustical transducers are positioned in the center of the group in question according to the arrangement topography of the orchestra, see the example in Fig. 35-3.

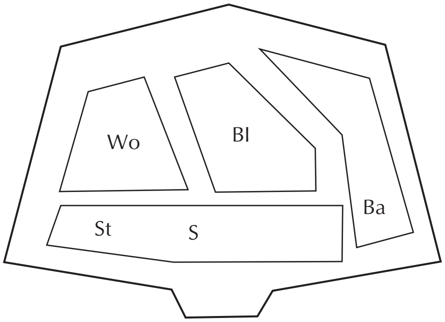


Figure 35-3. Typical arrangement of simulated orchestra-groups on a concert hall platform.

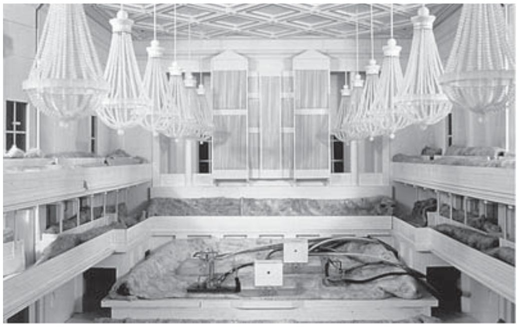
By evaluating the measured binaural room impulse responses on the basis of the register balance measure-(BR, see Section 7.2.2.15); it is possible to infer room-acoustical measures for a required architectural modification of the horizontal and vertical boundary surfaces of the performance zone, and to clarify questions concerning the vertical staggering of the orchestra. The sound intensity-time behavior allows conclusions regarding the sound attack of the individual instrumental groups, and regarding masking effects in the frequency and time domains from which measures for the acoustical formation of the secondary structure of the room can be inferred.

Fig. 35-4A shows an example for the use of the scaled-model measuring technique for a concert hall, as compared to the original room, Fig. 35-4B.

35.1.2 Building a Computer Model, Entering Room Data

Entering room data into a simulation program must be simple and straightforward. A combination of graphical and numerical entry of the data, planes, and vertex points has to be supported. Entering the room into the program must be efficient in order to make the program work cost effective and intuitive. If the room entry takes too long, the program becomes much less valuable as a real design tool. There are different ways to enter room data:

- By X, Y, Z coordinates.
- By text files.
- By import from professional drawing programs like AutoCAD or SketchUp.
- By use of drawing tools like prototypes or predefined room shapes.



A. Acoustical indoor room model scale 1:20 with simulated orchestra groups.

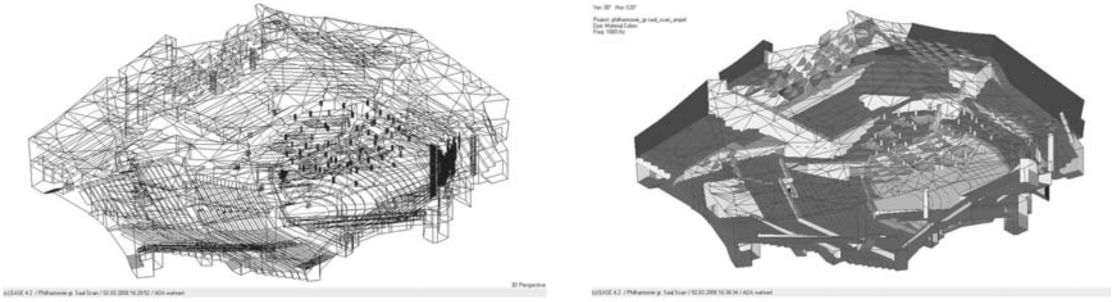
B. Finished original room.

Figure 35-4. Konzerthaus, Berlin.

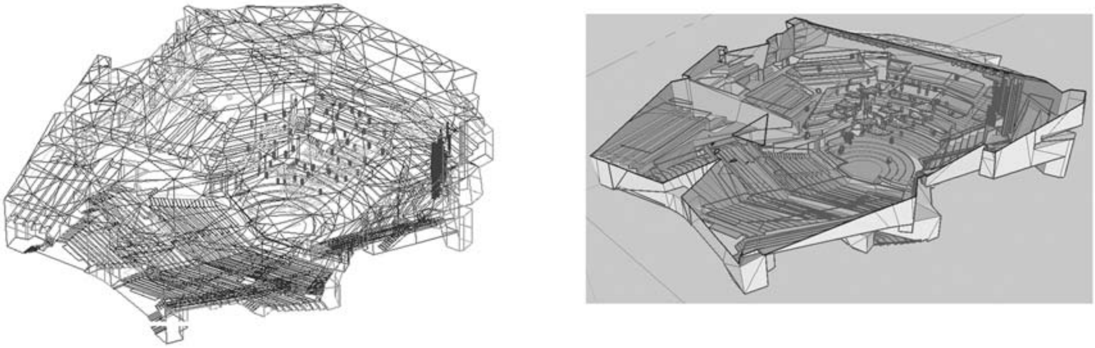
Fig. 35-5A to D shows one example of a model with different view options.

Simpler models that normally should have between 500 and a maximum of 1500 faces may be created based on simple room shapes or prototypes. A manipulation routine should allow one to stretch or shrink dimensions to adapt the prototype to the requirements. This way a simple room model for basic investigations can be created within minutes.

Better would be the possibility of importing DWG or DXF or other similar architecture files directly. But the disadvantage here is that architects in the early design phase do not create 3D models and offer only 2D drawings. These drawings are of less use and so the acoustician has to enter the model vertex by vertex, line by line and area by area. Sometimes 3D models can be built by expanding a 2D plan and by manipulating the result.



A. Wireframe model. B. Rendered view.



C. AutoCad presentation. D. SketchUp view.

Figure 35-5. Same room model in four different views.

35.1.3 Acoustic Sources and Loudspeaker Systems

35.1.3.1 Natural Sources

Sound reinforcement quite often has to deal with reinforcing natural sources like human voices or natural music instruments. Therefore to know the amount of reinforcing required, we have to know the quality of the reference sources and the following aspects:

- Loudness.
- Frequency response.
- Directivity.

Fig. 35-6 shows the level and frequency range of natural sources and instruments perceived by human beings. In this range natural sources develop their sound power and produce level components in the mentioned frequency range. Everything outside this range is on one side masked by noise (lower 30 dB in the midrange) or dangerous for our health—pain threshold ≈120 dB. Frequency components lower than 25 Hz are becoming inaudible as well as frequency parts higher than 15–20 kHz—depending on age and health.

Natural sound sources like human voices or musical instruments do not radiate sound in an omnidirectional way. It comes close in case of a human voice, but the higher the frequencies, the more the head becomes an obstacle to the sound radiation backward. Fig. 35-7A shows the directivity balloon curves of a female voice in the vertical domain. Below 1000 Hz the pattern is almost an omnidirectional radiation pattern, but for higher frequencies the radiation dominates more and more in front of the head. This is also the reason that in such concert halls where the audience is behind the orchestra, people sometimes complain about the singer's clarity.

The radiation behavior of musical instruments is much more complex. Here a lot of investigations have been done, especially by Meyer.[6] Fig. 35-7B shows the 3D presentation of the directivity balloon of a horn instrument including the player. According to Meyer the shadow effect of the player himself is also considered. We observe with increasing frequencies a reduced radiation into the front domain of the player.

To model all these different natural sources correctly, their radiation behavior has to be known and this not only as single instruments but also in groups. Here a lack of corresponding data is still evident.

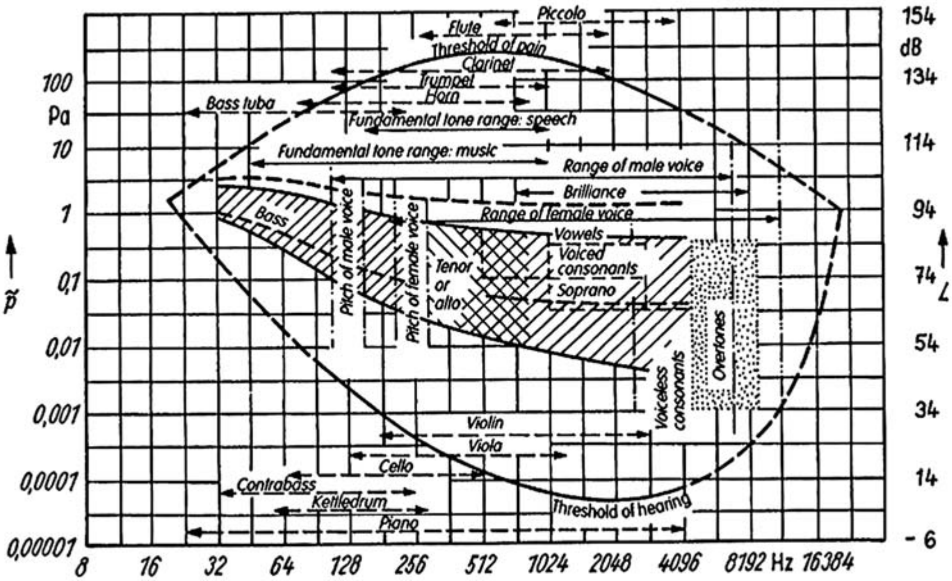


Figure 35-6. Level and frequency range of natural sources and instruments. Courtesy V. O. Knudsen 1932.[5]

35.1.3.2 Loudspeaker Types

We distinguish the following loudspeaker types:

- Point sources.
- Loudspeaker columns.
- Cluster.
- Line arrays.
- Digitally controlled sound columns.

For the use of these different sound radiators, their performance parameters must be known. One will soon note, however, that the performance parameters specified by the manufacturers vary in accuracy and scope. For nearly two decades the Standards Committee of the AES has been trying to update rules and standards for a uniform approach in this respect.[7] When studying the data sheets of diverse manufacturers one will nevertheless note considerable discrepancies allowing the expert to draw conclusions as to the quality of the data given. For this reason we are going to mention the most important data to be specified in loudspeaker design. Let us start with the so-called point sources.

35.1.3.2.1 Point Sources

Point sources do not show automatically omnidirectional radiation behavior. Their directivity behavior is measured on a turntable and all directivity balloon data is referred to the point of rotation, therefore the name point sources.

Transfer Behavior. The nominal *load capacity* P_n of this loudspeaker type is the rms electrical power specified by the manufacturer according to the design characteristics.

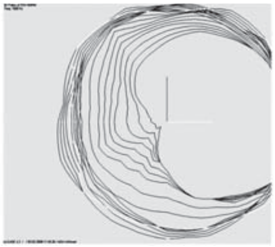
The ratio between the sound pressure $\tilde{p}$ and the voltage $\tilde{u}$ required to attain this capacity at the radiator is called *sensitivity* T_s.

$$T_s = \frac{\tilde{p}}{\tilde{u}} \qquad (35\text{-}3)$$

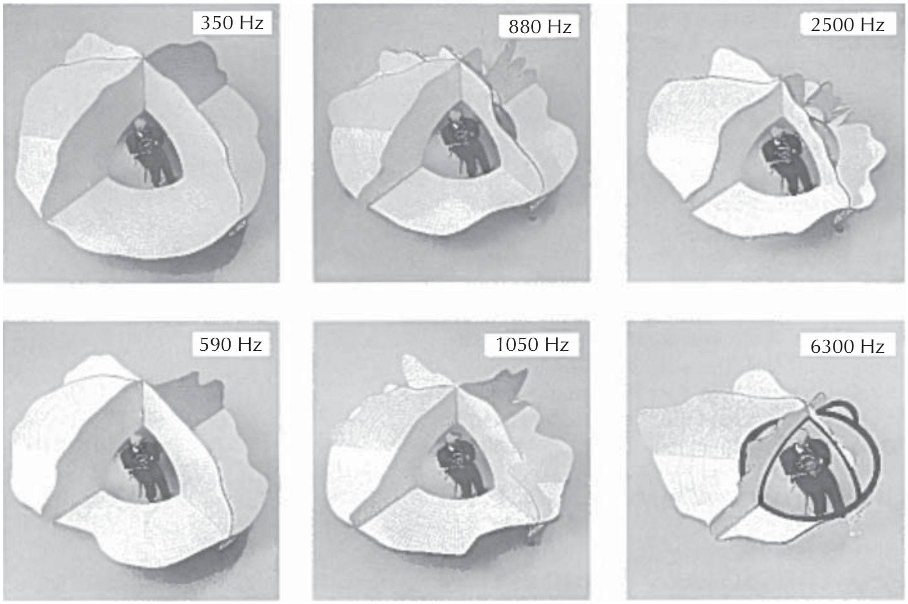
One distinguishes between a *free-field sensitivity* T_d and a *diffuse-field sensitivity* T_r. The free-field sensitivity is normally indicated for a reference point on the reference axis at a distance of 1 m from the loudspeaker. This can be expressed by

$$T_d = \frac{\tilde{p}_d\, r}{\tilde{u} r_0} \qquad (35\text{-}4)$$

The diffuse-field sensitivity has to be ascertained in a diffuse field, for instance in a reverberant chamber. In order to eliminate the room property characterized by the equivalent absorption area of the room, a correction factor has to be used:



A. Female voice.



B. Horn instrument including his player.

Figure 35-7. Directivity pattern of a natural source and an instrument.

$$T_s = \frac{\tilde{p}}{\tilde{u}} \tag{35-5}$$

where,

T_s and T_d or T_r are given in Pa/V.

The *sensitivity level,* G_S, is defined as the logarithmic quantity of sensitivity:

$$G_S = 20\log\frac{T_S}{T_0}\text{ dB}$$

The reference sensitivity T_0 is preferably 1 Pa/V. If another value is chosen, it has to be indicated.

The graphic representation of the sensitivity level as a function of frequency is called *frequency response*.

One of the quantities most frequently used in sound reinforcement engineering is the rated or characteristic sensitivity. In combination with the nominal load capacity it serves, among other things, to ascertain the maximum achievable sound pressure in the main reference axis of a loudspeaker or a loudspeaker system. According to the definition given in the Standard DIN 45570,[8] also AES2-1984 (r2003),[7] it is the ratio between the sound pressure p_d averaged over a determined frequency range (mostly 250–4000 Hz) and measured on the reference axis at a distance of 1 m from the reference point of the radiator—usually the point of rotation during measurements, and the square root of the power supplied. By the standard, this power is referred to the *nominal impedance* Z_n of the radiator as

$$P_s = (\tilde{u}^2/Z_n).$$

Thus the *rated sensitivity* is

$$\begin{aligned} E_K &= \frac{\tilde{p}_d}{\tilde{u}}\sqrt{Z_n\frac{r}{r_0}} \\ &= \frac{\tilde{p}_d\, r}{\sqrt{P_S}\, r_0} \end{aligned} \qquad (35\text{-}6)$$

where,
$\tilde{p}_d$ is direct sound pressure,
r is distance to the loudspeaker on the main axis,
r_0 is 1 m distance.

Because of its reference to power, this expression is also designated as rated power sensitivity. According to DIN 45570 T1, the logarithmic quantity of this expression is called *characteristic sound level* L_K, but also *sensitivity/*dB. It is defined by

$$\begin{aligned} L_K &= 20\log\frac{E_K}{E_{K0}}\ \text{dB} \\ &= 20\log E_K\ \text{dB} + 94\ \text{dB}. \end{aligned} \qquad (35\text{-}7)$$

An important parameter for approximating the sound-field conditions in rooms is the *front to random factor* γ. It characterizes the relationship between the acoustic power that would be radiated into the room by an omnidirectional loudspeaker having the same free-field sensitivity as the real loudspeaker to be assessed, and the acoustic power of the real loudspeaker:

$$\begin{aligned} \gamma &= \frac{\int_S \tilde{p}_0^{\,2} dS}{\int_S \tilde{p}^2(\vartheta) dS} \\ &= \frac{S}{\int_S \Gamma^2 dS} \end{aligned} \qquad (35\text{-}8)$$

where,
$\tilde{p}$ is sound pressure ($\tilde{p}_0$ is measured in the main front direction),
S is the globe surface around the loudspeaker,
ϑ is the room angle,
for Γ see Eq. 35-17.

A measuring procedure for ascertaining the *front to random factor* was established in the IEC publication 268-5 (1972):[9]

$$\gamma = \left(\frac{\tilde{p}_d\, r}{\tilde{p}_r r_H}\right)^2 \qquad (35\text{-}9)$$

where,
$\tilde{p}_d$ is direct sound pressure,
$\tilde{p}_r$ is reverberant sound pressure,
r is distance to the loudspeaker,
r_H is the critical distance in the diffuse sound field, see Eqs. 7-10.[10]

The *directivity factor* $Q(\vartheta)$ is often used for this term, but it is a function of the angle ϑ, see Eq. 35-20.

The logarithmic quantity of the *front to random factor* is the *front to random index*

$$C = 10\log\gamma\ \text{dB} \qquad (35\text{-}10)$$

It corresponds to the difference between the free-field and the diffuse-field sensitivity levels:

$$C = G_d - G_r \qquad (35\text{-}11)$$

where,
G_d is the sensitivity level in the direct field,
G_r is the sensitivity level in the diffuse field.

It is also expressed by the sound levels measured at 1 m distance in the direct field of the loudspeaker (L_d) and in the diffuse field (L_r) of a room having the reverberation time RT_{60} and the volume V,

$$C = L_d - L_r + 10\log\frac{RT_{60}}{V}\ \text{dB} + 25\ \text{dB} \qquad (35\text{-}12)$$

where,
L_d is the direct sound level,
L_r is the diffuse sound level,
RT_{60} is the reverberation time in seconds,
V is the volume of the room in m^3.

An equal input power P_{el} is taken for granted.

Because of the dimensions of the radiators, the wavelength of the radiated sound in the lower-frequency range is long compared to the radiating surface. Because of this difference there results only an insignificant directivity. With rising frequency the relationship changes and directivity increases.

For sound reinforcement purposes it has been proven in practice that slight increases of approximately 3 dB/octave of the front to random index of the loudspeaker system is appropriate, because most of the natural sound sources show a similar increase giving rise to a corresponding timbre change.

For approximate calculations or also measurements of the front to random index, one has to cover at least the range between 500 and 1000 Hz. Many manufacturers are indicating, in the data sheets of their products, the frequency dependence of directivity.

By means of the front to random index C and the nominal power rating P_n it is also possible to describe the characteristic sound level of a loudspeaker system:

$$L_K = L_W + C - 10\log P_n - 11\text{ dB} \tag{35-13}$$

where,
L_W is the sound power level.

The *efficiency* η of a loudspeaker system is determined by the ratio between radiated acoustic power and supplied electric power:

$$\begin{aligned}\eta &= \frac{P_{ak}}{P_{el}}\\ &= \left(\frac{E_K^2}{\rho_0 c} \times \frac{4\pi r_0^2}{\gamma_L}\right) \times 100\%\end{aligned} \tag{35-14}$$

where,
P_{ak} is the acoustic sound power,
P_{el} is the electrical power applied,
E_K is the sensitivity of the loudspeaker,
r_0 is 1 m distance,
γ is the front to random factor of the loudspeaker,
ρ_0 c is the characteristic acoustic impedance of air = 408 Pa s/m^3 at 20°C.

By combining all constants one obtains the following approximation:

$$\eta = 3\frac{E_K^2}{\gamma_L}\,\%\,. \tag{35-15}$$

This correlation can be seen in Fig. 35-8. The efficiency of loudspeaker systems lies in reality between 0.1% and 10%. As is the case with the rated sensitivity, the efficiency is often referred to the nominal impedance Z_n of the loudspeaker and designated as *nominal efficiency* η_n:

$$\eta_n = \frac{\tilde{p}_d^2 Z_n}{\gamma_L \tilde{u}^2} \times \frac{4\pi r^2}{\rho_0 c} \tag{35-16}$$

where,
$\tilde{p}_d$ is the direct sound pressure,
Z_n is the nominal impedance,
γ_L is the front to random factor of the loudspeaker.

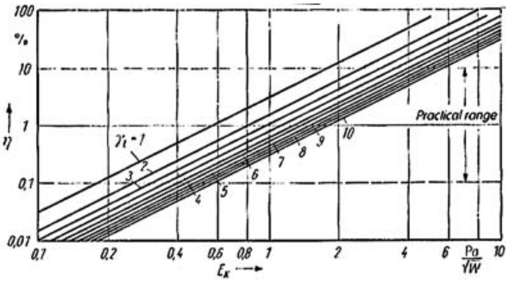


Figure 35-8. Efficiency of a loudspeaker as a function of rated sensitivity and front to random factor.

Eqs. 35-14, 35-15, and 35-16 suggest that because of the frequency dependence of the front to random index and the insignificant frequency dependence of the free-field sensitivity, the loudspeaker system efficiency may as well depend heavily on the frequency.

Directional Properties. All loudspeakers used in real life show a more or less pronounced directional dependence of radiation, which is frequency dependent—just like beaming behaviors. This angular dependence of sound radiation is characterized by three quantities that are going to be considered in detail.

The *angular directivity ratio* Γ for a frequency or a frequency band is the ratio between the sound pressure p radiated at an angle ϑ from the reference axis, and the sound pressure p_0 generated on the reference axis at equal distance from the selected acoustic reference point (this reference point is selected by the loudspeaker manufacturer and must be published in data sheets; generally it is the center of gravity of the loudspeaker box).[10]

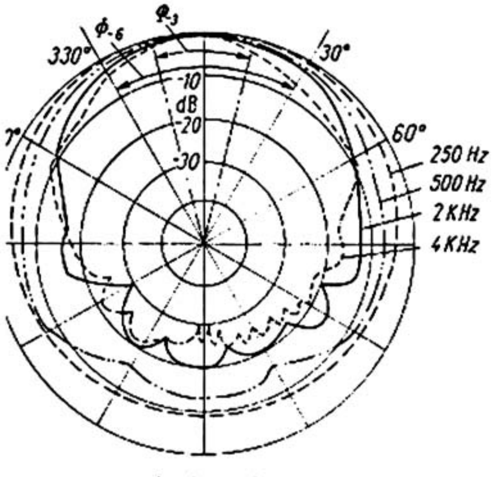
$$\Gamma(\vartheta) = \frac{\tilde{p}(\vartheta)}{\tilde{p}_0}\,. \tag{35-17}$$

In general $\Gamma(\vartheta) \le 1$. If the maximum of directional characteristics does not occur at $\vartheta = 0°$, then $\Gamma(\vartheta) > 1$.

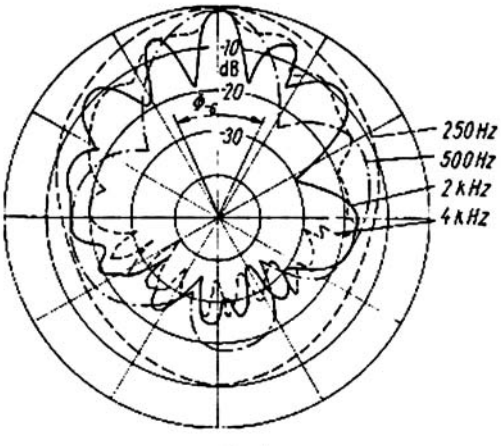
The logarithmic quantity of the angular directivity ratio is the angular directivity gain

$$D(\vartheta) = 20\log\Gamma(\vartheta)\text{ dB} \tag{35-18}$$

Fig. 35-9 shows the directional characteristic of the horn loudspeaker in a polar plot of the directivity gain. One sees the main maximum at 0° and several secondary maxima at higher frequencies.



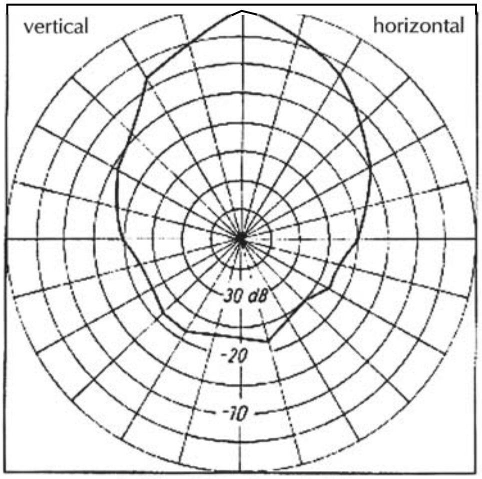
horizontal



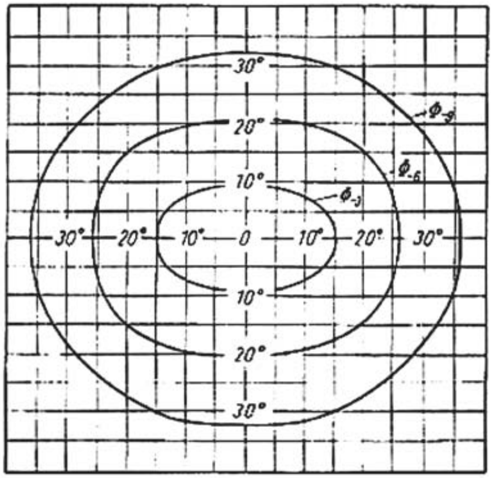
vertical

Figure 35-9. Polar plot of the angular directivity gain of a sound column with indication of the radiation angles.

if



A. Directivity characteristic.



B. Coverage area.

Figure 35-10. Plot of the directional effect of the radiator DML-1122 (Electro-Voice) frequency 2 kHz; front to random index 15 dB; maximum sound level at 1 m distance: 125 dB.

An important parameter for direct sound coverage is the angle of radiation Φ (beam width angle). It stands for the solid-angle margin within which the directivity gain drops by a maximum of 3 dB or 6 dB (or another value to be specified) as against the reference value. The curves of equal directivity gain are marked Φ_{-3}, Φ_{-6}, or generally Φ_{-n}; the higher the directivity the smaller the angle of radiation, Fig. 35-10.

Because of the curves of equal directivity gain and the sound distribution loss, the impact of direct sound of a loudspeaker on a surface may produce elliptic curves that represent a calculated SPL isobar area of the direct sound coverage. These isobar areas are important in the planning of sound reinforcement systems as coverage areas.

For combining the influence of the directional effect as well as that of the distribution between directional and omnidirectional energy, one uses, in acoustics, the *directivity deviation ratio*:[10]

$$\Gamma^{*}(\vartheta) = \sqrt{\gamma}\,\Gamma(\vartheta) \tag{35-19}$$

This quantity is also of high importance to sound reinforcement engineering. It characterizes the reverberation component caused by the loudspeaker in the excited room.

The square of this quantity is the well-known directivity factor Q:

$$\begin{aligned} Q &= g(\vartheta) \\ &= \gamma\Gamma^2(\vartheta) \end{aligned} \tag{35-20}$$

Especially in the United States, it is common to use just Q for different angles ϑ. Nevertheless this entity is angle-dependent and therefore it should always be referenced along with the corresponding angle. The logarithmic expression of the directivity factor $Q(\vartheta)$ is the so-called directivity index DI (also angle dependent)

$$\begin{aligned} DI &= \mathrm{H}(\vartheta) \\ &= 10\log g(\vartheta)\ \mathrm{dB} \\ &= 10\log Q(\vartheta)\ \mathrm{dB} \end{aligned} \tag{35-21}$$

where,

$H(\vartheta)$ is the reverberation directional index.

In the German literature one uses for the directivity factor Q, the reverberation directional value $g(\vartheta)$.[10] By the same token the directivity index, DI, is called *reverberation directional index* $H(\vartheta)$.

The reader should be aware of the partially contradicting conventions, of which some are using Q and DI only for values of $\vartheta = 0°$ and others employ Q and DI in an angle-dependent way, sometimes without clearly stating so.

Transmission Range. According to several standards, the transmission range of a loudspeaker is the frequency range usable or preferably used for sound transmission. That region of the transmission curve in which the level measured on the reference axis in the free field does not drop below a reference level generally characterizes the transmission range. The reference value is the average over the bandwidth of 1 octave in the region of highest sensitivity (or in a wider region as specified by the manufacturer). In the ascertainment of the upper and lower limits of the transmission range, there are not considered any peaks and dips whose interval is shorter than $^1/_8$ octave.

This definition implies that loudspeakers have to necessarily be checked as to their transmission range before being used in sound reinforcement systems. With radiators intended for indoor use, it is also necessary to consider the front to random factor—i.e., the influence of the diffuse-field component on the formation of the resulting sound pressure.

For special loudspeaker systems—e.g., studio monitoring equipment—narrower tolerance fields of the free-field sound pressure are indicated for the transmission range. Thus the OIRT Recommendation 55/1 permits for the range from 100 Hz–8 kHz a maximum deviation of ±4 dB from the average value, whereas below, down to 50 Hz, and above, up to 16 kHz, the tolerance field widens to −8 dB and +4 dB.[11]

Fig. 35-11 shows exemplary the behavior of free-field sensitivity, diffuse-field sensitivity, and front to random index of a radiator.

Moreover the transmission range is influenced, especially in the lower-frequency range, by the installation conditions or the arrangement of the radiator. Fig. 35-12 shows that the arrangement of the loudspeaker system has a considerable influence on the transmission curve. This is due to the fact that arranging the radiator in front of, below, or above a reflecting surface causes interferences of the direct sound by the strong reflections that give rise to comb-filter-like cancellations, which can be proven by a narrow-band analysis of the resulting signal. These cancellations are particularly pronounced, if the source is in front of a wall, and the radiator has compensating openings in its rear part, or if these reflections come from a distance of about 1.5 m out of a room corner—e.g., between ceiling and wall.

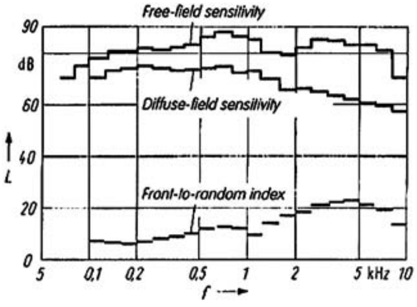


Figure 35-11. Frequency dependence of the front to random index, as compared with the free-field and the diffuse-field sensitivities.

As a rule one can say that the ear normally does not perceive dips and peaks that are not measurable in a ⅓ octave band filter analysis (unless they show pronounced periodic structures).

A good bass radiation is produced if the radiating plane is embedded in a reflecting surface, for instance, a wall or a ceiling. In this case there may also exist a certain angle between the radiating plane and the surrounding surface.

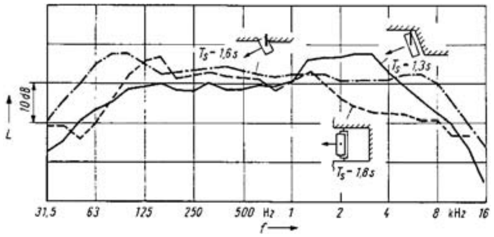
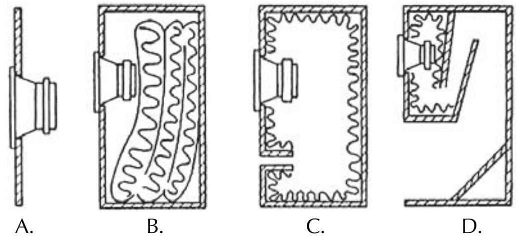


Figure 35-12. Frequency response curve of a loudspeaker at different mounting conditions. T_s is the reverberation time of the hall.

Types of Loudspeakers. The different tasks of sound reinforcement engineering require different radiator types. These differ as to size and shape of their enclosures, the form of sound conduction, the types of driving systems used, as well as arrangement and combination of the same. In this way one obtains different directional characteristics of sound radiation, sound concentrations, sensitivities, transmission ranges, and dimensions that facilitate solutions for diverse applications or even enable them at all.

Among the simplest radiators are single loudspeakers of smaller dimensions and ratings that are used in decentralized information systems, for instance, for covering large flat rooms or for producing room effects in multipurpose halls. The integration into a wall or an enclosure of these loudspeakers avoids the acoustic short-circuit usually seen with no baffle situations—suppressing the pressure compensation between the front and rear sides of the diaphragm. To this effect a baffle panel or an open or closed box may be used, Fig. 35-13.



A. Baffle panel.
B. Closed box.
C. Bass reflect box.
D. Transmission box.

Figure 35-13. Different measures for suppressing the acoustic short circuit.

With a closed box one has to consider that the oscillating part of the loudspeaker functions in one direction against the relatively stiff air cushion of the box. Loudspeakers for such *compact boxes* are for this reason provided with an especially soft diaphragm suspension so that they cannot be easily used for other purposes.

Acoustically more favorable are the conditions with vented enclosures, the *bass reflex boxes* or *phase reversal boxes*. Such box loudspeakers are nowadays used less as decentralized broadband radiators, but increasingly for high-power large-size loudspeaker arrays.

Another possibility for achieving a determined directional characteristic consists in the arrangement of sound-conducting surfaces in front of the driving loudspeaker system. Given that such arrangements are mostly of hornlike design, they are named *horn loudspeakers*. Because of the high characteristic sensitivity and the high-directional characteristic, this radiator design is very well suited for sound reinforcement in big auditoriums where the desirable frequency range and different target areas (coverage areas) require the use of different types of radiation patterns.

For technical reasons it is not sensible to construct a broadband horn for the overall transmission range. A better solution is several horn loudspeakers complementing each other.

Bass Horns. Owing to the great dimensions involved, the design of bass horns requires extensive compromises. Practical models of bass horns receive a horn shape, as a rule, only in one dimension, whereas at a right angle to it, sound control is achieved by means of parallel surfaces. The power-handling capacity of such bass horns, which are mainly used in music or concert systems, is about 100–500 VA.

Medium-Frequency Horns. The greatest variety of driver and horn designs is available for horn loudspeakers for the medium-frequency range of about 300 Hz–3 kHz.

The drivers used are mostly dynamic pressure-chamber systems connected to the horn proper by means of a throat, the so-called *throat-adapter*.

Treble Horns. For the upper frequency range, two main types of horn loudspeakers are produced. These are the horn radiators showing similar design characteristics as the medium-frequency horns that function in the frequency range from 1–10 kHz, and the special treble loudspeakers (*calotte horns*) used for the frequency range from 3–16 kHz.

35.1.3.3 Loudspeaker Line, Sound Column and Line Arrays with In-Line Arrangement of Radiators

Classical Columns. For many tasks of sound reinforcement engineering, one requires radiators capable of producing a high sound level at a large distance from their point of installation, while minimally affecting microphones located at close range to them. To have this effect they must show a determined directional characteristic and beaming. A radiator type suitable for this purpose is the loudspeaker array consisting—in the variant required—of a stacked arrangement of in-phase-identical loudspeakers. In the plane orthogonal to this arrangement, there occurs a pressure addition, whereas in the areas above and below this plane, there is a cancellation by interference because of the early-to-late difference between the components stemming from the different loudspeakers, Fig. 35-14. Each of the individual loudspeakers radiates the sound spherically and the sound waves get favorably superposed in the far field, whereas the effect of the individual loudspeaker prevails in the near field. For the far field the following equation was given by Stenzel[12,13] and Olson[14] for the angular directivity ratio Γ, the so-called polars.

$$\Gamma = \left| \frac{\sin\left[\frac{n\pi d}{\lambda}\sin a\right]}{n\sin\left[\frac{\pi d}{\lambda}\sin a\right]} \right| \tag{35-22}$$

where,

n is the number of individual loudspeakers,

d is the spacing of the individual loudspeakers,

a is the radiation angle,

λ is the wavelength of sound,

l is $(n-1)\,d$, which is the length of the loudspeaker line.

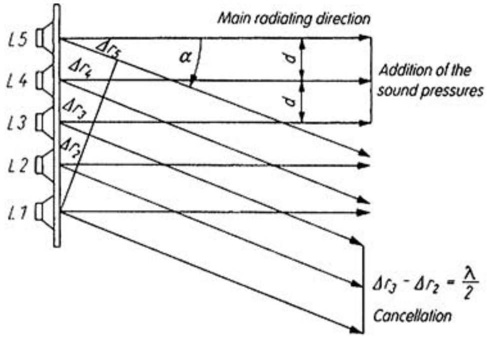

L5
L4
L3
L2
L1
Main radiating direction
Δr5
Δr4
Δr3
Δr2
α
d
d
Addition of the sound pressures
Δr3 - Δr2 = λ/2
Cancellation

Figure 35-14. Operating principle of a classical sound column.

This directional effect of a loudspeaker line according to Fig. 35-15[15] is shown in Fig. 35-16A—a balloon at 1 kHz, and in Fig. 35-16B—a balloon at 2 kHz. The line consists of nondirectional loudspeakers arranged with a spacing of 25 cm. Secondary maxima occur at frequencies above a critical frequency (wavelength = spacing of the loudspeakers), that is above 1400 Hz in the example. Thus a desirable disc-shaped radiation without secondary maxima can be observed at 1000 Hz, whereas at 2000 Hz lobes (secondary maxima) are already utterly evident.

The drawback of an in-line loudspeaker arrangement consists of the fact that

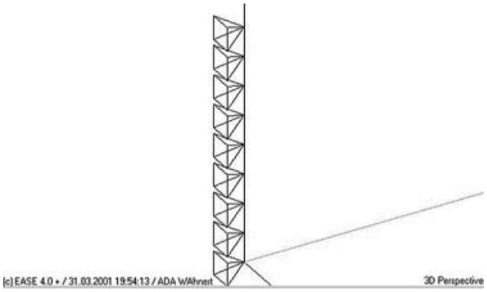
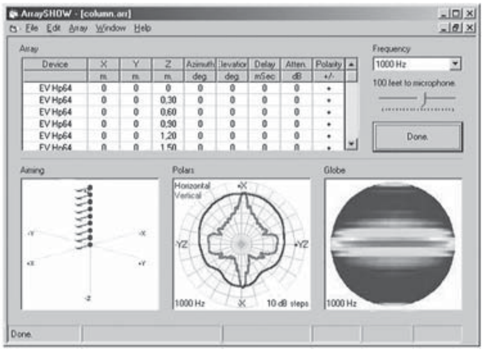

(c) EASE 4.0 + / 31.03.2001 19:54:13 / ADA WAhnert
3D Perspective

Figure 35-15. A line presentation with nine horns HP64 in the simulation program EASE.

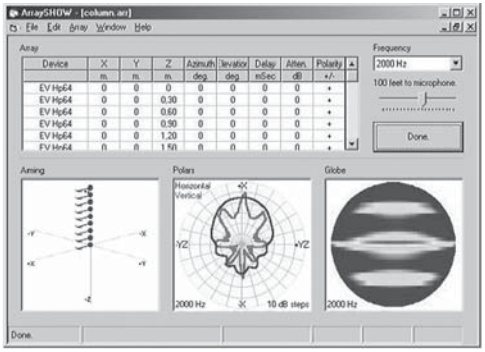

ArraySHOW - [column.arr]
File
Edit
Array
Window
Help
Array
Aiming
Polars
Globe
Horizontal
Vertical
+X
-X
-YZ
+YZ
-Y
+Y
+X
-X
-Z
1000 Hz
10 dB steps
1000 Hz
Done.

Frequency: 1000 Hz

100 feet to microphone

Device	X	Y	Z	Azimuth	Elevation	Delay	Atten.	Polarity
	m.	m.	m.	deg	deg	mSec	dB	+/-
EV Hp64	0	0	0	0	0	0	0	+
EV Hp64	0	0	0,30	0	0	0	0	+
EV Hp64	0	0	0,60	0	0	0	0	+
EV Hp64	0	0	0,90	0	0	0	0	+
EV Hp64	0	0	1,20	0	0	0	0	+
EV Hp64	0	0	1,50	0	0	0	0	+

A. 1000 Hz.


ArraySHOW - [column.arr]
File
Edit
Array
Window
Help
Array
Aiming
Polars
Globe
Horizontal
Vertical
+X
-X
-YZ
+YZ
-Y
+Y
+X
-X
-Z
2000 Hz
10 dB steps
2000 Hz
Done.

Frequency: 2000 Hz

100 feet to microphone

Device	X	Y	Z	Azimuth	Elevation	Delay	Atten.	Polarity
	m.	m.	m.	deg	deg	mSec	dB	+/-
EV Hp64	0	0	0	0	0	0	0	+
EV Hp64	0	0	0,30	0	0	0	0	+
EV Hp64	0	0	0,60	0	0	0	0	+
EV Hp64	0	0	0,90	0	0	0	0	+
EV Hp64	0	0	1,20	0	0	0	0	+
EV Hp64	0	0	1,50	0	0	0	0	+

B. 2000 Hz.

Figure 35-16. Balloon presentation of the line array according Fig. 35-15 in a simulation.

- The desired directional effect is given only in a range below the critical frequency, whereas above that frequency there occur additional secondary maxima.
- The directivity is frequency-dependent—*front-to-random factor* γ of the main maximum $\approx 5.8\ lf$ (l is the length of the column in m and f the frequency in kHz).[16]
- The directivity increase does not only occur in the directivity domain, but also, owing to the distances of the individual loudspeakers, in the scattering domain, so the column is losing directivity at high frequencies.

All these frequency-dependent properties of the loudspeaker lines involve the possibility for timbre changes to occur over the width and depth of the covered auditory. In order to eliminate or limit this drawback, the lines are often subdivided in the upper frequency range. This is mostly accomplished by curving the line "bananalike" or like a so-called J-Array. Alternatively individual elements can be rotated slightly off-axis in the horizontal domain, such as in alternating angles of ±10 degrees relative to the aiming axis of the system.

Line Arrays. Modern line arrays do not consist of a line up of individual cone loudspeakers, but instead of a linear arrangement of wave-guides of the length l, which produce a so-called coherent wave front. In contrast to the traditional sound columns, these arrays radiate in their near range so-called cylindrical waves. This near range is frequency dependent and only valid up to the following distances r:

$$r = \frac{l^2}{2\lambda} \qquad (35\text{-}23)$$

where,

array length and wavelength are in meters.

In 1992 Christian Heil was the first to present this new design at the AES in Vienna.[17] With the product V-DOSC by L-Acoustics, a new technology was introduced which can now be found with modifications in the product range of more than forty manufacturers (compare Fig. 35-17).

The characteristic feature of these systems is that the sound levels decrease in the near range by only 3 dB with distance doubling, and begin to decrease like those of spherical radiators only beyond the near range. This way it is possible to cover large distances with high sound levels and without having to use delay towers.

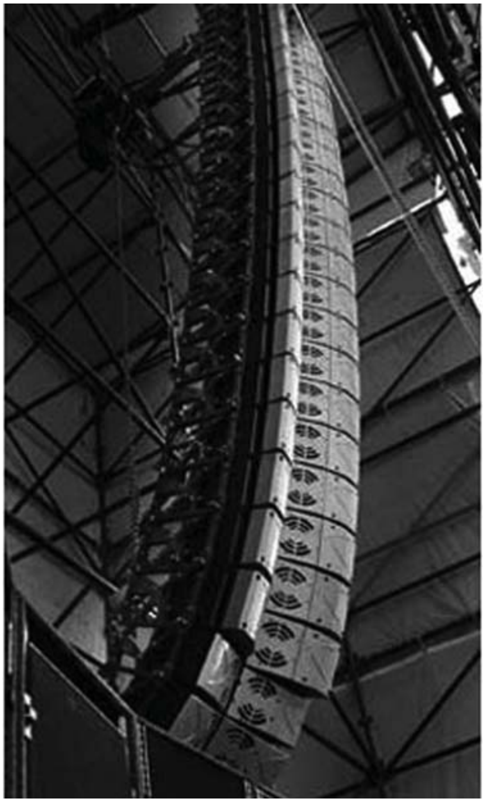


Figure 35-17. A Geo T series column by NEXO SA.

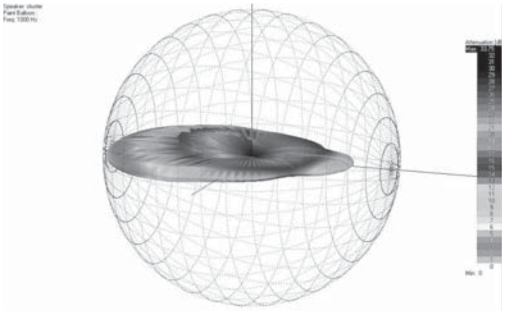
Digitally Controlled Line Arrays. A way of reducing the frequency dependence of the directional characteristics and beaming of sound lines consists of supplying the sound signal with different phases and levels to the individual loudspeakers in an array.

Duran-Audio was one of the first manufacturers that reduced the length of its Intellivox loudspeaker lines with increasing frequency by electronic means (so-called DDC solution). This solution resulted in loudspeaker lines with pronounced directivity in the vertical domain and constant sound power concentration in the horizontal domain.[18] Fig. 35-18 illustrates such a directional effect in 3D representation.

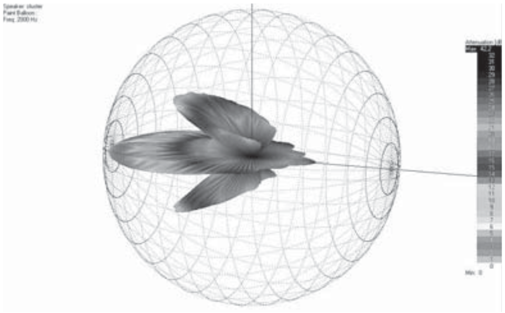
Other manufacturers go similar ways, Renkus-Heinz with the ICONYX loudspeaker,[19] Fig. 35-19, the French company ATEIS (Messenger),[20] and EAW (DSA series).[21]

By changing the firmware control the following features of such columns are possible:

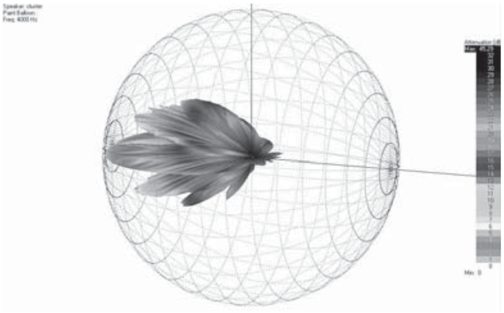
1. Constant SPL versus distance.
   - Midband frequencies.
   - Noncomplex shaped audience areas.
2. The performance is optimized with the following parameters:



A. 1000 Hz.



B. 2000 Hz.



C. 4000 Hz.

Figure 35-18. Cluster balloon presentation of Intellivox 2C in EASE.

- Opening angle.
- Aiming angle.
- Focus distance.
- Mounting height with respect to audience area.

This solution, however, finds its limitations with a complicated audience-area layout. Moreover, certain sound level distributions can be obtained only after several corrections. This gives rise to questions like: "How can a loudspeaker array be controlled so as to create a predefined far and near field response?" One approach was made by Duran Audio with the digital directivity synthesis (DDS).[22]

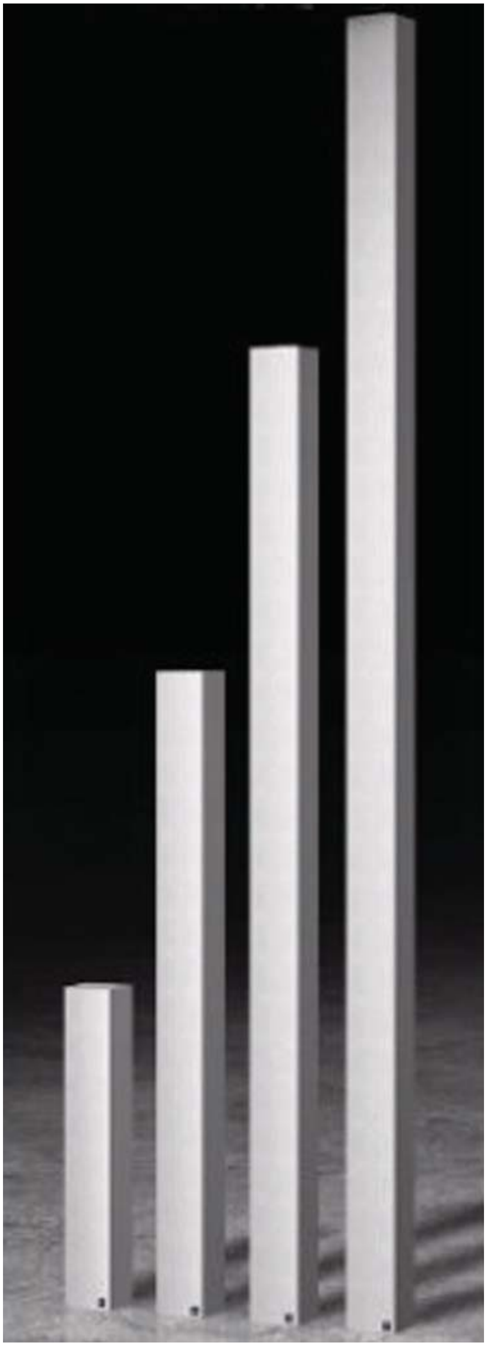


Figure 35-19. ICONYX Series from Renkus-Heinz Inc.

Here the directivity pattern of an array is adaptable to audience areas; and uniform SPL distribution also becomes possible in complex-shaped audience areas.

It stands to reason that point sources or line loudspeakers of a certain dimension cannot exhibit this radiation behavior automatically. Here the manufacturer has to supply along with the arrays, not only the electronic driver unit for the same, but also the parameter setup algorithm. By means of attached software this algorithm will then be controlled according to the desired application.

35.1.4 Wall Materials

To simulate the radiation behavior of sources in rooms or open spaces we need to construct corresponding models. All boundary walls of these models need to have the corresponding acoustic properties:

- Absorption.
- Scattering.
- Diffraction.

These properties have been discussed in Section 7.3.4, therefore they will not be repeated here. Instead, some specialties important to know when doing computer modeling should be added here.

35.1.4.1 Absorber Data

Absorbing behavior has been known for hundreds of years, but data has been available only for 80 years. We distinguish between data measured in a reverberation chamber (Standard ISO 354)[23] or data that is angle dependent. The latter absorption coefficients are very seldom measured and only available for special applications. For computer simulation, the absorption coefficient measured in the diffuse field will be used. This coefficient is measured by the corresponding manufacturers and published in specification brochures. It is measured in octave or ⅓-octave bands and starts normally at 63 Hz and goes up to 12 and even 16 kHz. In most simulation programs the low end is skipped because the actual simulation routines do not cover frequency ranges below 100 Hz. The highest-frequency band is quite often only 8 kHz.

All this data is meanwhile published in table form and some simulation programs have more than 2000 materials from different manufacturers on board.

35.1.4.2 Scattering Data

Scattering data is not found in textbooks except for some special scattering materials or samples. Here should be mentioned the products of RPG Diffuser Systems Inc., which produce special modules with sound-scattering surfaces.

On the other hand it is known that the absolute value of the scattering coefficient s is less important. The fact is that there is almost no material not scattering ($s = 0$) or only scattering ($s = 1$). The practical values for the scattering coefficients are between 0 and 1. So there are some rules of thumb to define the actual scattering coefficient in simulation software programs. Some programs give some guidance to estimate the coefficients, other programs like EASE 4.2 use special BEM routines (compare to Fig. 7-46) to derive the coefficient in a way as it should be measured according the proposals of Mommertz,[24] see also Standard ISO 17497-1.[25]

A scattering coefficient will never generally be available in tables (except the mentioned special module values), because the way the interior architect uses the materials in a hall affects the scattering behavior.

Therefore the scattering behavior of wall parts in a computer model must be determined model specific.

35.1.4.3 Diffraction, Low-Frequency Absorption

As we will see in Section 35.3.2, the computer simulation programs use different ray-tracing algorithms to calculate the impulse responses in model rooms. But these routines of using particle radiation are only valid above a certain frequency determined by

$$f_1 = K\sqrt{\frac{RT_{60}}{V}} \qquad (35\text{-}24)$$

where,

K is a constant (2000 in metric units, 11,885 in U.S. units,

RT_{60} is the apparent reverberation time in seconds,

V is the volume of the room in cubic meters or cubic feet.

For lower frequencies and especially in small rooms the particle assumption cannot be applied. Here the wave acoustics routines are applied. An analytical solution is impossible, so numeric routines have been developed. Mainly the finite element method (FEM) and the boundary element method (BEM) are used. First for applying the FEM, the computer model must be subdivided in small volumes (meshes), where the dimensions of the mesh correspond with the upper frequency handled by the FEM. The higher the frequency, the smaller the dimensions and the longer the calculation time. As an example, to build a mesh in a hall of 10,000 m^3 you need a mesh resolution of about 280,000 subvolumes to apply the FEM up to 500 Hz. Fig. 35-20 shows such a mesh grid for a church model. For the BEM only the surface must be meshed accordingly.

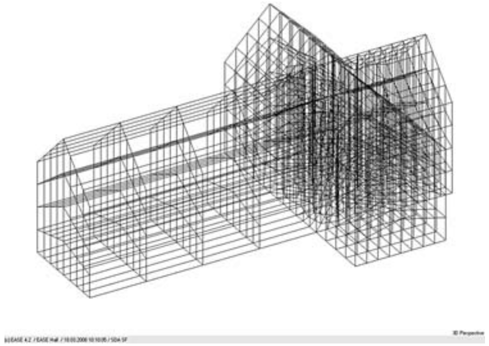


Figure 35-20. Meshed model in EASE 4.2.

After the mesh is ready and that is a quite difficult job in complex room structures,[26] we need to know the impedance behavior of the single wall parts. This is again quite complex because any stiffness or mass values of the majority of the wall materials are not known. So in a first approach the impedance of the wall material can be derived from the known absorption coefficient. Now by applying the well-known algorithm of the FEM, the transfer function at selected receiver places may be calculated. By means of a Fourier transformation you obtain the impulse response in the time domain. By means of this method, transfer functions at receiver places may also be calculated, even if the receiver is shadowed from the sending position and the direct sound was only coming by diffraction to the receiver.

This method can be used very well in small rooms below 300 Hz. A mesh of a control room of 135 m^3 consists only of around 1000 subvolumes, if frequencies higher than 300 Hz are neglected. This way very fast calculation results can be expected.

35.1.5 Receivers and Microphone Systems

35.1.5.1 Human Ears

The properties of the human ears are explained in a lot of books about psychoacoustics, including Chapter 3 of this handbook. In simulation programs the acoustic properties of a room or the free-field environment are determined by calculation of the so-called impulse response. This response is calculated using ray-tracing methods. For a single point in space, the so-called monaural response is determined, and the result supplies not only the level at the receiver place, but also the frequency dependence, the angle of incidence for single reflections, and the run-time delay in comparison to the first incoming signal (direct sound). Using so-called head-related transfer functions (HRTF) measured with dummy heads, or using in-ear microphones, Fig. 35-21,[27] the monaural impulse response may be converted into a binaural one used for real-time convolution, see Section 35.3.3.

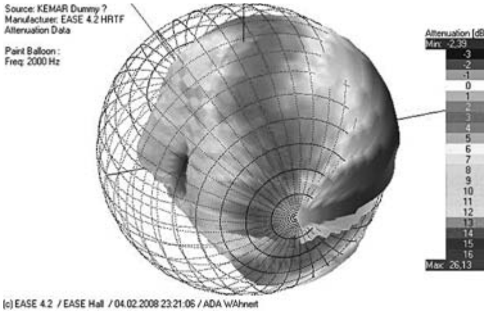


Figure 35-21. HRTF balloon of the left ear of a dummy head.

35.1.5.2 Microphones

The use of microphones in sound reinforcement systems requires observation of a number of conditions. To avoid positive acoustic feedback, it is frequently necessary to keep the microphone at closer distance to the sound source so that often considerably more microphones have to be used. Moreover the live conditions demand very robust microphones.

To simulate the use of microphones, to precalculate the acoustic feedback threshold or to simulate enhancement systems based on electronic processing, an exact knowledge of the properties of the microphone types and their connection technique is needed.

Basic Parameters. The microphone data are laid down in standards.[28] In this context we will consider only this data that is important for computer modeling. For further information especially, regarding the types of microphones, please refer to Chapter 16.

The magnitude of the output voltage of a microphone as a function of the incident sound pressure is described by the microphone sensitivity

$$T_E = \frac{\tilde{u}}{\tilde{\tilde{p}}} \qquad (35\text{-}25)$$

in V/Pa or its 20-fold common logarithm, the *sensitivity level*

$$G_s = 20\log\frac{T_E}{T_0}\text{ dB} \qquad (35\text{-}26)$$

The reference sensitivity T_0 is normally specified for 1 V/Pa.

Depending on the test conditions one distinguishes the following sensitivities:

- The *pressure open-circuit sensitivity* T_{Ep} as the ratio between the effective output voltage at a certain frequency and the effective sound pressure of a vertically incident sound wave.
- The *free-field open-circuit sensitivity* T_{Ef} which considers by special measuring conditions the pressure increase conditioned by the cross-section dimensions of the microphone.
- The *diffuse-field sensitivity* T_{Er}, which reflects the diffuse sound incident on the microphone.

Directivity Behavior. The dependence of the microphone voltage on the direction of incidence of the exciting sound is called *directional effect*. The following quantities are used for describing this effect:

- The *angular directivity ratio* $\Gamma(\vartheta)$ is the ratio between the free-field sensitivity T_{Ed} for a plane sound wave arriving under the angle ϑ to the main microphone axis and the value ascertained in the reference level (incidence angle 0°).

$$\Gamma(\vartheta) = \frac{T_{Ed}(\vartheta)}{T_{Ed}(0)} \qquad (35\text{-}27)$$

- The *angular directivity gain D* is the twenty fold common logarithm of the angular directivity ratio.
- The *coverage angle* is the angular range within which the directivity gain does not drop by more than 3 dB, 6 dB, or 9 dB against the reference axis.

Apart from the quantities describing the ratio between the sensitivities of the microphone with sound incidence from various directions deviating from the main axis, it is also necessary to describe the relationship between the sensitivities with reception of a plane wave and those with diffuse excitation. With these quantities it is then possible to ascertain the suppression of the room-sound components against the direct sound of a source to be transmitted. This energy ratio is described by the following parameters:

- The front to random factor is the ratio between the electric power rendered by the microphone when excited by a plane wave from the direction of the main axis, and the power rendered by the microphone excited in a diffuse field with the same sound level and same exciting signal. If the sensitivity was measured in the direct field as T_{Ed} and in the diffuse field as T_{Er}, the front to random factor results as

$$\gamma_M = \frac{T_{Ed}^{\,2}}{T_{Er}^{\,2}}. \qquad (35\text{-}28)$$

- The front to random index is the ten fold common logarithm of the front to random factor.

While the front to random factor of an ideal omnidirectional microphone is 1, that of an ideal cardioid microphone is 3. This means that a cardioid microphone picks up only ⅓ of the sound power of a room picked up by a comparable omnidirectional microphone at the same distance from the source. This implies, for instance, that with identical proportion of the sound power, the speaking distance for a cardioid microphone may be three times greater than that of an omnidirectional microphone.

35.2 Transducer Data for Acoustic Simulation

To simulate an entire acoustic system, all parts must be taken into account. Besides the room, loudspeakers and natural sound sources as well as microphones and the human hearing system have to be considered. In this section, our main goal is to review existing practice and outline advantages and disadvantages that the user of a software program should be aware of when applying performance data for a particular sound transducer. In this regard, our intention is to talk about the simulation of transducers with respect to the electroacoustic and room-acoustic prediction of the acoustic system as a whole. We will not be concerned with mathematical methods applied in the design process of loudspeakers or microphones. These usually provide a much higher degree of accuracy in some regards, but at the same time often provide insufficient data for other simulation purposes. Specifically, for transducer design utilizing BEM/FEM-based prediction methods we refer the reader to available textbooks and publications.

35.2.1 Simulation of Loudspeakers

In computer aided acoustic design and especially for sound reinforcement applications, the level of accuracy to which sound sources are modeled plays a crucial role. Accordingly, most simulation software packages have continuously developed their capabilities of describing loudspeakers by measurement data along with the complexity of the loudspeaker systems themselves. At the same pace, the availability and fast development of personal computers had a significant impact on acoustic measurement systems and their accuracy on the one hand and on the computing power available on the other hand.

In this sense, the measurement and simulation of loudspeaker systems can be roughly divided into two periods of time. The first period, until the late 1990s, was characterized by the use of simplified far-field data for almost any sort of loudspeaker and the assumption of a point-source-like behavior. But with the advent of modern line array technology, for both tour sound and speech transmission applications, new concepts had to be developed. These methods include the use of multiple point sources as well as advanced mathematical models to image the complexity of today's loudspeaker systems. In addition to that, research was further accelerated by the broad availability of DSP platforms and the resulting need to simulate DSP controlled loudspeakers as well as the virtual disappearance of computer-based constraints, such as calculation speed and memory.

35.2.1.1 Simulation of Point Sources

Theoretical Background. For many years the radiation behavior of sound sources, and loudspeakers in particular, was basically described by a 3D matrix containing magnitude data in a fixed spectral and spatial resolution. Starting with the late 1980s, typical data files contained directivity data for the audible octave bands, such as from 63 Hz to 8 kHz, and for a spherical grid with an angular spacing of 15°. Mostly, data was also assumed to be symmetric in one or two planes. With the need for higher data resolution and the limits of available PC memory and computing power changing at the same time, more advanced data formats developed eventually reaching a nowadays typical resolution of 5° angular increments for ⅓-octave frequency bands. Tables 35-1 and 35-2 show some of these typical loudspeaker data formats and their resolutions.

Now let us look at the background for this development. We express the complex sound pressure $\bar{p}$ for the time-independent propagation of a spherical wave:[29]

$$\bar{p}_{Sphere}(\vec{r},f) = \frac{\bar{A}(\varphi,\vartheta,f)}{|\vec{r}|}\exp(-j\vec{k}\vec{r}) \qquad (35\text{-}29)$$

where

$\vec{r}$ is the receiver location,

f is the frequency,

$\bar{A}(\varphi,\vartheta,f)$ is the complex radiation function of the source depending on angles φ and ϑ (both being functions of $\vec{r}/|\vec{r}|$) as well as on the frequency,

$\vec{k}$ is the wave vector.

Loudspeaker measurements $\hat{A}$ happen at discrete angles φ_k, ϑ_l, and frequencies f_m, the simulation software has to interpolate between such data points to obtain a smooth response function:

$$\bar{p}_{Sim}(\vec{r},f_m) = \frac{f_{Int}(\hat{A}(\varphi_k,\vartheta_l,f_m))}{|\vec{r}|}\exp(-j\vec{k}\vec{r}) \qquad (35\text{-}29A)$$

Here the interpolation function is represented by f_{Int}. The frequency resolution is basically given by the set of available data points of f_m, the angular resolution is given by the density of data points φ_k, ϑ_l.

For a long time, most measurements were made to acquire magnitude data only, $\hat{A} = |\hat{A}|$ and $f_{Int}((\hat{A})) = |f_{Int}(\hat{A})|$. In such a case, the simulation of interaction between multiple sound sources yields a sound intensity, I_{Sum}, that is derived either by power summation for incoherent sources n (located at $\vec{r}_n$):

$$\begin{aligned} I_{Sum}(\vec{r},f_m) &= \left|\sum_n \bar{p}_n\right|^2 \\ &= \sum_n \frac{[f_{Int}(\hat{A}_n(\varphi_k,\vartheta_l,f_m))]^2}{|\vec{r}-\vec{r}_n|^2} \end{aligned} \qquad (35\text{-}30)$$

or by at least considering the run time phase $\Phi_n = -\vec{k}_n(\vec{r}-\vec{r}_n)$ for coherent sources:

$$I_{Sum}(\vec{r},f_m) = \left|\sum_n \bar{p}_n\right|^2$$

$$= \left|\left[\sum_n \frac{f_{Int}(\hat{A}_n(\varphi_k,\vartheta_l,f_m))}{|\vec{r}-\vec{r}_n|}\right]\exp[-j\vec{k}_n(\vec{r}-\vec{r}_n)]\right|^2$$

(35-30A)

This simulation model makes some significant assumptions:

1. First, the use of a spherical waveform assumes that both measurement and simulation happen in the far field of the device, that is, at a distance where the sound source (normally a surface) can be considered as a point source $\bar{p}_{Real}(\vec{r},f) \approx \bar{p}_{Sphere}(\vec{r},f)$.

Table 35-1. Conventional Loudspeaker Data Formats and EASE GLL Format

Data Type	EASE SPK	EASE XHN	GDF	ULYSSES UNF	CLF	EASE GLL
	Simple Data Table	Simple Data Table	Simple Data Table	Simple Data Table	Simple Data Table	Advanced Description Language
Balloon Symmetries	Full, Half	Full	Full, Half, Quarter	Full, Half, Quarter	Full, Half, Quarter	Full, Half, Quarter
Angular Resolution	5°	5°	5° or 10°	5° or 10°	5° or 10°	1° to 90°
Frequency Resolution	⅓ Octave	⅓ Octave	¹⁄₁ or ⅓ Octave	¹⁄₁ or ⅓ Octave	¹⁄₁ or ⅓ Octave	Any
Complex Data	Yes	Yes	No	No	No	Yes
Individual Transducers	No	No	No	No	No	Yes
Filters	No	No	No	No	No	Yes
Configurable	No	No	No	No	No	Yes

Table 35-2. Measurement Parameters for Typical Balloon Resolutions

Measuring Resolution	Measuring Points	Measuring Time (for 10 seconds per measuring point)	Implemented in
2 measuring planes, 15°, symmetrical in both measuring planes	24	≈ 10 min	EASE 1.0
Measurement on sphere surface, 10°, symmetrical in the horizontal plane	325	≈ 1 h	EASE 2.1 ULYSSES 1.0/CLF
Measurement on sphere surface, 10°, no symmetry assumptions	614	≈ 1½ h	CATT-Acoustic/CLF ODEON BOSE Modeler
Measurement on sphere surface, 5°, symmetrical in the horizontal plane	1297	≈ 3½ h	CLF EASE 3.0-4.2 EASE 4.2 DLL/GLL
Measurement on sphere surface, 5°, no symmetry assumptions	2522	≈ 7 h	ULYSSES 2.82/CLF EASE 3.0-4.2 EASE 4.2 DLL/GLL
Measurement on sphere surface, 2°, no symmetry assumptions	16022	≈ 2 d	MAPP (Meyer) EASE 4.2 DLL/GLL
2 Transducer measurements on sphere surface, 10°, no symmetry assumptions	1228	3 h	EASE 4.2 GLL EASE 4.2 DLL

2. Second, it is assumed that the density of discrete data points is high enough and the frequency and angular dependency of the directivity characteristics smooth enough so that the true radiation function of the spherical wave can be approximated by $\bar{A} \approx f_{Int}(\hat{A})$.
3. Third, the use of magnitude-only data, assuming that $\bar{A} \approx f_{Int}(\hat{A})$, requires that the point of reference during the measurement of $\hat{A}$ is chosen in a way that the true run-time phase, otherwise included with the measurements, can be reconstructed by the run-time phase $\vec{k}\vec{r}$ in the model. It requires that the source-inherent phase is negligible as well, $\arg\bar{A} \approx 0$.
4. It is assumed that the concerned loudspeaker system is a fixed system that cannot be changed by the user or when its configuration is changed its performance data is not affected. The measurement data is regarded as representative for the whole range of possible applications and configurations.
5. Finally, for the use of such point sources in computations involving geometrical shadowing and ray-tracing calculations, the source is regarded as located at a single point and is either wholly visible (audible) for a receiver or not.

These assumptions have been made especially in the early 1990s in order to obtain and use loudspeaker directivity data in a practical manner. Important factors were the availability and accuracy of measurement platforms and methods, the storage size of the processed measurement data and the PC performance with regard to processor speed available to the average user of the data.

However, these assumptions have a set of drawbacks. That became most evident with the broad use of large-format touring line arrays and digitally controlled loudspeaker columns but also with the increasing use of inexpensive DSP technology employed for multiway loudspeaker systems. Some of the issues conflicting with above points 1–5 are listed in the following.

- A large line array system of some meters, height cannot be measured adequately in its far field in addition to the fact that a large number of line array applications actually happens to take place mainly in the near field. Therefore the simulation of a whole line array as a point source is not valid within reasonable error ranges.
- Another problem often encountered is insufficient angular resolution. Loudspeaker columns and multiway loudspeakers exhibit significant lobing behavior in the frequency ranges where multiple acoustic sources interact at similar strength. Often too coarse angular measurements fail to capture these fine structures and thus cause erroneous simulation results due to aliasing/sampling errors.
- While in many cases the phase of the sound pressure radiated by a simple loudspeaker is negligible at least if it is considered on-axis and the run-time phase is compensated for, the same is not true for most real-world systems. On the one hand, for multiway systems one cannot generally define a single point where the measured phase response vanishes for all frequencies and angular directions. This is the problem of the so-called acoustic center for a set of sound sources. In such cases the measured phase data will typically show a run-time phase component that depends on angle and frequency, no matter where the point of rotation is. On the other hand, the inherent phase response plays an important role in describing the radiation behavior that is influenced by diffraction about the edges of the loudspeaker case, that is, at angles of 60 degrees and more off-axis.
- Loudspeaker systems become increasingly configurable, so that the user can adapt them to a particular application. Typical examples include almost all touring line arrays where the directional behavior is defined mechanically by the splay angles between adjacent cabinets, and in loudspeaker columns or multiway loudspeakers, where the radiation characteristics can be changed electronically by manipulating the filter settings.
- In advanced computer simulations of sound reinforcement systems in venues, geometrical calculations must be performed. This is required to obtain exact knowledge of which part of the audience might be shadowed by obstacles between the sound sources and the receivers. Geometric considerations are also needed in ray-tracing calculation in order to find reflections and echoes. For both processes, the reduction of a physically large loudspeaker system to a point source can lead to significant errors. Depending on the choice of the reference point for the source, particular reflections might not be found or are exaggerated, or a large fraction of the audience area might be seemingly shadowed by a very small object.

In addition to the above, a set of minor problems is also evident. This includes the definition of maximum power handling capabilities of multi-input systems that are represented by a single point source. The availability of case drawings to help in the mechanical design and the clear indication of the reference point that was used for the measurements are is important.

As a result of the obvious contradictions, a variety of proposed solutions emerged in the later 1990s. This development happened partially by the loudspeaker manufacturers and partially by the creators of simulation software as well. To resolve the problem of large-format loudspeaker systems, a subdivision into smaller elements is required to be able to measure them and use them for prediction purposes. To properly model the coherent interaction between these elements, complex measurement data, including both magnitude and phase data, is needed.

The most prominent solutions can be summarized as follows. Instead of measuring a whole system, so-called far-field cluster balloons were calculated based on the far-field measurement of individual cabinets or groups of loudspeakers.[30] To describe individual sound sources, phase data was introduced in addition to the magnitude-only balloon data.[31,32] Mathematical models providing phase information implicitly were applied, such as minimum phase or elementary wave approaches as well as 2D sound sources.[17] However, these first approaches lacked generality and thus their implementation into existing simulation software packages was specific, difficult, or even impossible.

The situation was resolved first by the concept of the loudspeaker DLL (dynamic link library), which serves essentially as a programmable plug-in for simulation software.[33] Another concept, namely the GLL (generic loudspeaker library), introduced a new loudspeaker data file format that is significantly more flexible than the conventional data formats, and is designed to resolve most of their apparent contradictions.[34] We will review both approaches in the next section as they have turned out to be a standardized way to model complex loudspeaker systems.

Practical Considerations. Improvements that are theoretically desirable must also be practically accomplished. It is clear that only reasonable measurement times can provide reliable data in an efficient manner. In practice, an angular resolution of 5 degrees has proven to be adequate for most needs, sometimes even lower resolutions of 10 degrees can be sufficient. Simulation software packages should be able to handle higher resolutions as well, but only for special cases. This is particularly feasibl when measuring durations can be reduced using multiple microphones, like ten or nineteen receivers arranged on an arc. This technique requires some care in the measurement setup since all of the microphones have to be calibrated and normalized relative to each other.

Also the acquisition of both magnitude and phase data requires more care than just the measurement of magnitude-only data. However, modern impulse response acquisition platforms provide a good means to obtain complex data and a sufficient frequency resolution. The representation of a loudspeaker directivity function based on impulse response wave files is thoroughly discussed by the Working Group of the Standards Committee SC04-01 of the AES.[7] As we will present farther below, the utilization of phase data in acoustic modeling has become an important factor. As an illustration, Fig. 35-22A—D show, the magnitude and phase data for a loudspeaker—UPL1 from Meyer Sound Inc.—in high resolution in both MATLAB and EASE.[35]

Additionally, it is worth mentioning—and we will give some practical guidelines in the next sections—that, in order to obtain acceptable data for a point source approach, measurements have to take place in the far field of that assumed point source. Like indicated previously, this may be difficult for large multiway cabinets or column loudspeakers.

In general it must be emphasized that the computer model utilizing this loudspeaker data can only be as good as the data of the lowest quality included. Nowadays, the accuracy of the loudspeaker data is often much higher than that of the material data. Absorption and scattering coefficients are usually only known in $^1\!/\!_1$ or $^1\!/\!_3$ octave bands for random incidence. The user must be aware that although loudspeaker direct field predictions may be very precise, any modeling of the reflections and the diffuse sound field in the room will be limited by the available material data. Furthermore, it is not very likely that there will ever be systematic, large-scale measurements of angle-dependent complex directivity data for the reflection and scattering of sound by wall materials.

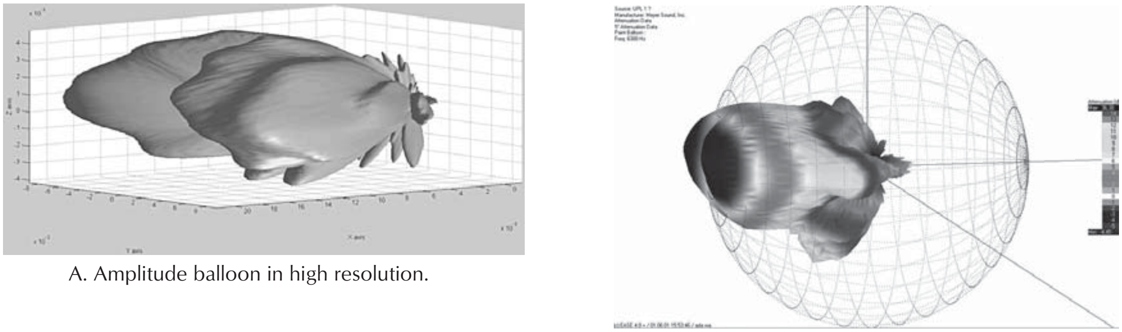
To complete this practical perspective another point of concern has to be underlined. Any data set describing the acoustic characteristics of a loudspeaker should also document important measurement parameters and conditions. In particular, the point of rotation used for the sensitivity and balloon measurements must be defined in such a data set and indicated in the case drawing as well, Fig. 35-23**.** Only when this reference point is known will the end user be able to define precisely the location of the loudspeaker in the computer model and to obtain the right results.

35.2.1.2 Simulation of Complex Loudspeakers

35.2.1.2.1 Modeling by Means of a DLL Program Module

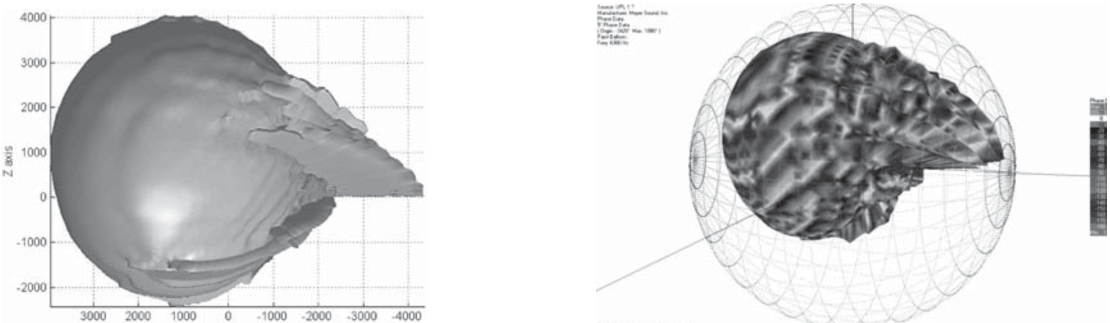
In a first step to overcome the variety of issues related to the reduction of complex loudspeakers to simple point sources, the DLL approach was developed.[36,.37] Technically speaking, the MS Windows dynamic link library (DLL) is a program or a set of functions that can be executed and return results. It cannot be run standalone but only as a plug-in of another software that accesses it through a predefined interface. The basic idea is to move the complexity of describing a sound source from the acoustic simulation program into a separate module that can be developed independently and that can contain proprietary contents. In this way, a clear cut is made between the creators for simulation software packages and the loudspeaker manufacturers who can develop product-specific DLL modules on their own.

However, acoustic prediction programs have different underlying concepts and therefore, although the DLL concept is a general philosophy, the DLL interfaces are different too. In consequence, a DLL built for one simulation platform cannot be used for another. Nevertheless, all DLL models share a similar approach to resolve the given problems. Because they can be programmed, they are essentially able to handle any kind of data and realize any kind of algorithm. If the mathematical description and/or sufficient measurement data for the loudspeaker system exists, this information can be encoded into a DLL. Given an appropriate theory for how the source radiates sound, the solution can be implemented without much compromise. Practically, the DLL provides the data describing the radiation of sound by a particular loudspeaker and the simulation software employs this data to model the interaction of the source with the room. For an example see Fig. 35-24.



A. Amplitude balloon in high resolution.

B. Amplitude balloon in EASE 4.0 resolution.



C. Phase balloon in high resolution.

D. Phase balloon in EASE 4.0 resolution.

Figure 35-22. High resolution and EASE 4.0 amplitude and phase presentation.

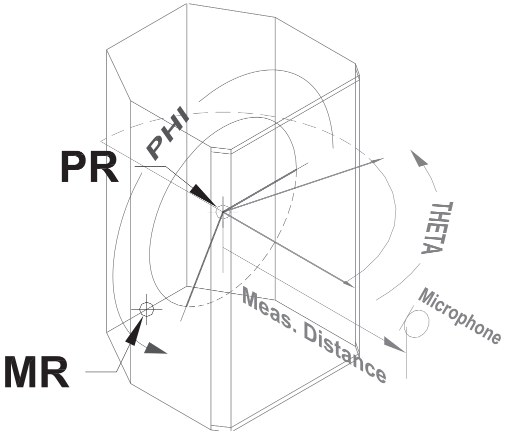
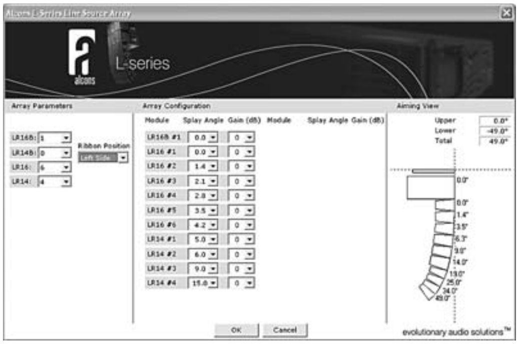


Figure 35-23. AutoCAD drawing of a loudspeaker showing the point of rotation PR and the case (mechanical) reference point MR.


Alcons L-Series Line Source Array
L-series
alcons
evolutionary audio solutions™
OK
Cancel

Array Parameters

LR16B: 1
LR14B: 0
LR16: 6
LR14: 4
Ribbon Position: Left Side

Array Configuration

Module	Splay Angle	Gain (dB)
LR16B #1	0.0	0
LR16 #1	0.0	0
LR16 #2	1.4	0
LR16 #3	2.1	0
LR16 #4	2.8	0
LR16 #5	3.5	0
LR16 #6	4.2	0
LR14 #1	5.0	0
LR14 #2	6.0	0
LR14 #3	9.0	0
LR14 #4	15.0	0

Aiming View

Upper 0.0°
Lower -49.0°
Total 49.0°

Figure 35-24. Screenshot of an EASE DLL.

Compared to conventional, tabular data formats this flexibility is unequaled. It is obvious that with a mathematical model of sufficient accuracy many of the previously discussed issues can be resolved. But at the same time, the development of a DLL module requires some programming effort. As an encoded binary file it is also proprietary to the loudspeaker manufacturer, so normally the user of simulation software and DLL plug-ins cannot determine how the loudspeaker system is actually modeled. Unless sufficient information is published by the creator of the DLL, the end user cannot estimate the level of prediction accuracy the model provides.

35.2.1.2.2 Modeling by Means of a GLL Data File

While trying to solve the same problems, technically the GLL concept[34] takes a different path compared to the DLL philosophy. Based on the experience with many loudspeaker manufacturing companies and the implementation of simulation and measurement software packages, the generic loudspeaker library (GLL) was developed as an object-oriented description language to define the acoustic, mechanical, and electronic properties of loudspeaker systems, Table 35-1. Since for each physical entity the GLL language has a representation in the software domain, there is no need to make artificial assumptions in order to comply with rigid, reduced data structures. Basically, in the GLL philosophy every sound radiating object should be modeled as such and every interaction possible between engineer and loudspeaker in the real world should be imaged in the software domain. In this picture, transducers, filters, cabinets, rigging structures and a whole array or cluster are present in the GLL with their essential properties and parameters, Fig. 35-25.

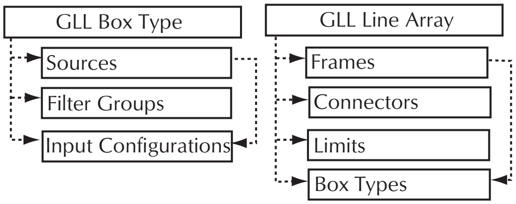


Figure 35-25. Some elementary objects of the GLL description language.

Typically, the GLL model of a loudspeaker consists of one or multiple sound sources, each with its own location, orientation, directivity, and sensitivity data. These sources can be simple point sources but also spatially extended sources, such as lines, pistons, etc. In addition to that, a complex directivity balloon based on high-resolution impulse response or complex frequency response data on a spherical grid describes the radiation behavior. With the sources representing the acoustic outputs of the loudspeaker on the one hand, the builder of a GLL defines the electronic inputs of the loudspeaker on the other hand. A filter matrix provides the logic to combine inputs and outputs, see the example in Fig. 35-26.

It can include multiple sets of filters, including IIR and FIR filters, crossover, and equalization filters. The loudspeaker box is mechanically characterized by means of a case drawing and data for center of mass calculations. Boxes can be combined into arrays and clusters. Available configurations are predetermined by the loudspeaker manufacturer according to the functions available to the end user. Additional mechanical elements such as frames and connectors allow specifying exactly which configuration possibilities exist.

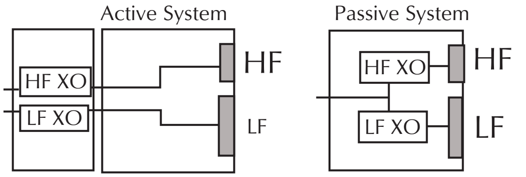


Figure 35-26. Typical GLL input configurations and filters for a two-way system.

Once all the data is assembled, the GLL is compiled into a locked, distributable file, Fig. 35-27. In fact, the end user of a compiled GLL can only see the loudspeaker system as he would see it in the real world. The user can apply filters to the electronic inputs of the loudspeaker and he can calculate (= measure) the acoustic output of the loudspeaker. He can look at the loudspeaker case as well. When modeling arrays he may change the arrangement of boxes as allowed by the manufacturer.

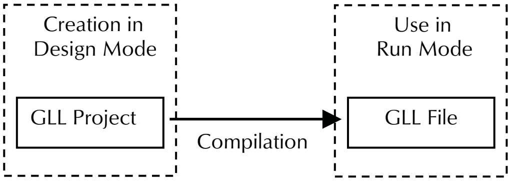


Figure 35-27. Compilation to create a GLL.

It is obvious that the GLL format provides a natural, straightforward way to describe loudspeaker systems. By means of a GLL model any active and passive multiway loudspeakers, digitally controlled column loudspeakers, line arrays, or loudspeaker clusters can be accurately represented with regard to their acoustic, electronic, and mechanical properties. Nonetheless the GLL model will fail due to its very nature, when artificial algorithms are to be implemented that do not have a counterpart in the real world.

35.2.1.3 Background of Simulation and Measurement

This section reviews simulation methods and measurement requirements as well as their theoretical basis with respect to both DLL and GLL modeling approaches.

35.2.1.3.1 Resolving the Far-Field Problem

One of the primary points to address in the simulation of the acoustic sources is the correct application of the data for the near field and far field. In the previous section we have emphasized that many loudspeakers and loudspeaker systems employed in the field today are actually used mainly in their near field, that is, at distances where the system cannot be approximated by a single point source with a distance-independent directivity pattern. Because of their size these systems can hardly be measured as a whole in their far field. However, measurements at a near-field distance are only valid for use at that distance and not beyond, see Eq. 35-23.

Principally, there are two solutions to that. On the one hand, one can try to model the system as what it is, namely a spatially extended source. It could be characterized mathematically by an ideal straight or curved line source with some correction factors derived from measurements. On the other hand, already for the purposes of practical assembly, transport, and maintenance, almost all large-format loudspeakers are composed of individual elements. For example, a touring line array is built of multiple cabinets each of which in turn house multiple transducers. Thus it seems natural that the line array is described primarily by its components and its overall radiation characteristics are derived from that. In consequence, representing the significantly smaller elements individually as point sources, now the measurement and the simulation only have to happen in the far field of the respective element. Coherent superposition of the sound waves radiated by these elements will then yield the correct behavior of the entire system for both near and far field.

35.2.1.3.2 Acquisition and Interpolation of Complex Data

Data. The issue of using complex data instead of magnitude-only data is closely related to finding an accurate way to interpolate data points over angle and frequency properly for both magnitude and phase data. In return, using complex data on the level of individual elements eliminates the need for higher precision when measuring and interpolating data on the level of the loudspeaker system as a whole.

Critical Frequency.

First, let us review the error that we make when measuring a loudspeaker directivity balloon about a given point of rotation (POR). Problems usually arise from the fact that one or several sources are slightly off-set from the POR and thus the measured data suffers a systematic error. For a given setup, Fig. 35-28, we can estimate the relative error for the magnitude data $\Delta|\hat{A}|$ for large measuring distances.[38]

$$\vartheta|\hat{A}| = 1 - \frac{x^2}{2d^2} + \frac{x}{d}\sin\vartheta \qquad (35\text{-}31)$$

where,

x is the distance between the POR and the concerned acoustic source,

d is the measurement distance between the POR and the microphone (with $d >> x$),

ϑ is the measurement angle between the microphone and the loudspeaker axis.

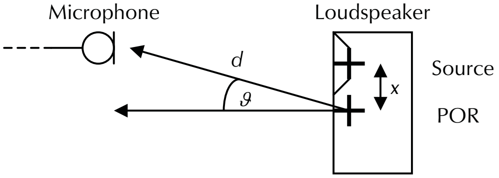


Figure 35-28. Typical setup for loudspeaker measurements.

The error is maximal for measurements where the connecting line between microphone and POR passes through both POR and acoustic source, in this case, at an angle of $\vartheta = 90°$. Nevertheless, for all practical cases the error is largely negligible. For example, typical values of $x = 0.1$ m and $d = 4$ m yield an error of only 0.2 dB.

When describing loudspeakers by magnitude data only, the phase is neglected completely. To simulate the interaction between coherent sources, often the run-time phase calculated from the time of flight between POR and receiver is used. As stated earlier, this assumes that the inherent phase response of the system is negligible and that there is an approximate so-called acoustic center where the run-time phase vanishes and which must be used as the POR. For this measurement situation the systematic error in the phase data, $\delta\Phi = \delta \arg A$,can be calculated as well.[38] For large measuring distances d, see Fig. 35-28, it is given by

$$\delta\Phi = \frac{2\pi}{\lambda} x|\sin\vartheta| \qquad (35\text{-}32)$$

for magnitude-only data $(\arg\hat{A} \approx 0)$

where,

λ denotes the wavelength.

In contrast, by acquiring phase data in addition to the magnitude data, this offset error can be minimized:

$$\delta\Phi = \frac{2\pi x^2}{\lambda 2d}(1-\sin^2\vartheta) \qquad (35\text{-}33)$$

for complex data ($\arg\hat{A} \neq 0$)

Hence, by using phase data the error is reduced by an order of magnitude. In practice, it is most useful to define a maximum acceptable phase error, such as

$$\delta\Phi_{Crit} = \frac{\pi}{4}$$

and use that to derive an upper (critical) frequency limit based on the measurement setup. Fig. 35-29 shows this critical frequency f_{Crit} as a function of the distance between POR and acoustic source.

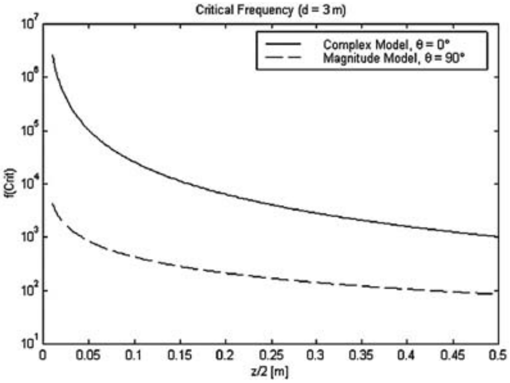


Figure 35-29. Critical frequency for magnitude data and complex data as a function of the distance x = z/2 between POR and acoustic source, at a measuring distance of 3 m and a maximal phase error of 45°.

We emphasize that the use of phase data does not only reduce the error in the directivity data but it also largely eliminates the need to define, find, and use the so-called acoustic center, the imaginary origin of the far-field spherical wave front.

Local Phase Interpolation. Once complex directivity data is available for a loudspeaker, the next step is to define an appropriate interpolation function for the discrete set of data points to image the continuous sound field of the source in the real-world ($f_{Int}(\hat{A}) \rightarrow \bar{A}$). An algorithm will have to work for both magnitude and phase data in both domains, frequency and space. While averaging, smoothing, and interpolating magnitude data is usually a straightforward task, the same is not true for phase data. Due to the mathematical nature of phase, its data values are located on a circle. Accordingly, when phase is mapped to a linear scale the interpolation has to take wrapping into account. In this regard, a variety of methods have been proposed, such as use of group delay, unwrapped phase, or the so-called continuous phase representation. Although these methods have their advantages, it could be shown that generally none of them is appropriate for full-sphere radiation data of a real loudspeaker.[39]

Alternatively, a method named local phase interpolation can be applied successfully when some care is taken about the resolution of the underlying data. This method essentially interpolates phase on a local scale rather than globally. For example, let the phase average of two data points i and j be defined as

$$\langle\Phi\rangle = \frac{1}{2}\Phi_i + \frac{1}{2}\Phi_j \qquad (35\text{-}34)$$

Then, it is assumed that the corresponding phase data points are all located within a certain range:

$$|\Phi_i - \Phi_j| < \frac{\pi}{2} \qquad (35\text{-}35)$$

In this respect, i and j may represent two angular data points ϑ_i and ϑ_j or two frequencies f_1 and f_2. Also, the averaging or interpolation function may involve more than two points.

Note that in the above case we have assumed that for calculating the absolute difference the maximum possible difference is π. This can always be accomplished by shifting the involved phase values by multiples of 2π relative to each other.

From the condition above we can derive requirements directly for the measurement. Assuming that the phase response will be usually dominated by a run-time phase component due to one or several acoustic sources being located away from the POR, conditions for the spatial and spectral density of data points can be computed.[39] With respect to frequency one obtains

$$x_{crit} \approx \frac{c}{4\Delta f} \qquad (35\text{-}36)$$

where,

Δf denotes the frequency resolution,

c the speed of sound.

Given these parameters, x_{crit} is the maximal distance allowed between the POR and the acoustic source at the given frequency resolution. With regard to angle, one finds analogously

$$x_{crit} \approx \frac{c}{4f\sin(\Delta\vartheta)} \qquad (35\text{-}37)$$

where,

f is the frequency,
$\Delta\vartheta$ is the angular resolution.

As an example, these limits correspond roughly to a measurement setup where the acoustic source is not farther away than ca. 0.15 m from the POR. Phase data points will be close enough up to a frequency of 8 kHz, if the frequency resolution is at least 1⁄12 octave (or 475 Hz) and the angular resolution is at least 5 degrees. Such conditions are well within what is possible with modern measurement platforms.

Data Acquisition. Although loudspeaker performance data and directivity patterns have been measured for several decades, no definitive standard has emerged from that practice. Also for some years now, the AES standards committees try to unify the variety of existing methods and concepts to reach some commonly accepted measuring recommendations.

The accurate measurement of loudspeaker polar data is one of the issues of the ongoing discussion. Especially the acquisition of complex frequency response data, which asks for significantly higher accuracy in the measurement setup, and better control of the environment than the measurement of magnitude-only data. To determine the exact phase response of the loudspeaker under test relative to the POR, it is inevitable to measure and compensate the measuring distance as well as the environmental conditions that influence the propagation of sound along that path. For example, to be exact within a quarter of a wavelength at 8 kHz, all distance measurements must be accurate within less than a centimeter of length. Although this is not a trivial task, professional acoustic laboratories have been built at the factories of manufacturers, at universities, and by independent service providers.[40] As a result, today many loudspeaker systems are measured using measurement platforms that can provide high-resolution impulse response or complex frequency response data.

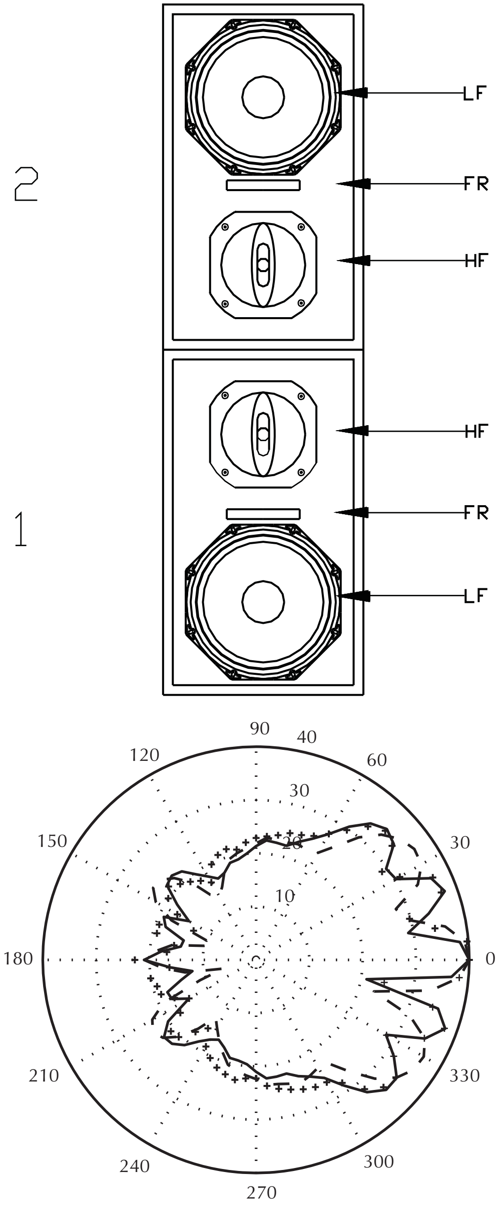
But it is important to note that gathering measurement data as described above only slightly increases the overall effort. To build a measurement setup capable of acquiring complex balloon data means a high initial effort, but with respect to automated polar measurements, the subsequent measuring durations are the same as for magnitude-only data. The measurement of the individual components of a loudspeaker cabinet or array is obviously connected with longer measurement times. However, in many cases the angular resolution for a transducer measurement may be lower than for the full multiway device because its directivity behavior is much smoother. In the same manner, the frequency resolution can be chosen adequately. Finally, the acquisition of phase data also means that the so-called acoustic center does not have to be determined in a time-consuming procedure. Mounting the loudspeaker for a measurement is therefore much simpler. The measurement of different transducers of the same loudspeaker does not require remounting the device anymore, as well. Additionally, as we will show below, loudspeaker designers and manufacturers gain direct benefits from advanced measurement data, such as directivity prediction, crossover design, and verification capabilities.

Figs. 35-30, 35-31, and 35-32 show some of the advantages gained by using complex data for individual components. Fig. 35-30 shows a comparison of measurement versus prediction for a stacked configuration of two two-way loudspeakers, arranged horn to horn (HF-HF). Its vertical directivity pattern at 1 kHz is displayed in Fig. 35-30, measured data (+ curve) and calculation based on complex data (solid curve) are in good agreement. Calculations with magnitude-only data (dashed curve), provide erroneous results. In this case, the port of the loudspeaker (FR) was chosen to be the POR. A similar discrepancy between measurement and prediction using magnitude-only data can be seen in the arrangement of woofer to woofer (LF–LF), Fig. 35-31. To illustrate the sampling problems described before, Fig. 35-32 shows the same configuration at 4 kHz. Here measurements (+ curve) can only be imaged properly by a computation at angular increments of 2.5°. The dashed curve is using individual components measured at 5°. Computations or measurements at a too coarse resolution of 5° (dashed curve) fail completely to describe the properties of the system when being interpolated.

Due to the complexity of establishing an accurate and phase-stable measurement setup, a set of alternative approaches is practiced. This includes, in particular, the modeling of the wave front radiated at the loudspeaker by elementary point sources according to the Huygens principle. Other models are based on the idea of deriving the missing phase response from the magnitude response, such as by the minimum phase assumption. Some of these implementations work quite well for a subset of applications, such as in the vertical domain or within some opening angle relative to the loudspeaker's axis. But generally these ideal models lack the means to depict the sound radiating properties of the loudspeaker in those domains where it is not so well behaved and analytically treatable.

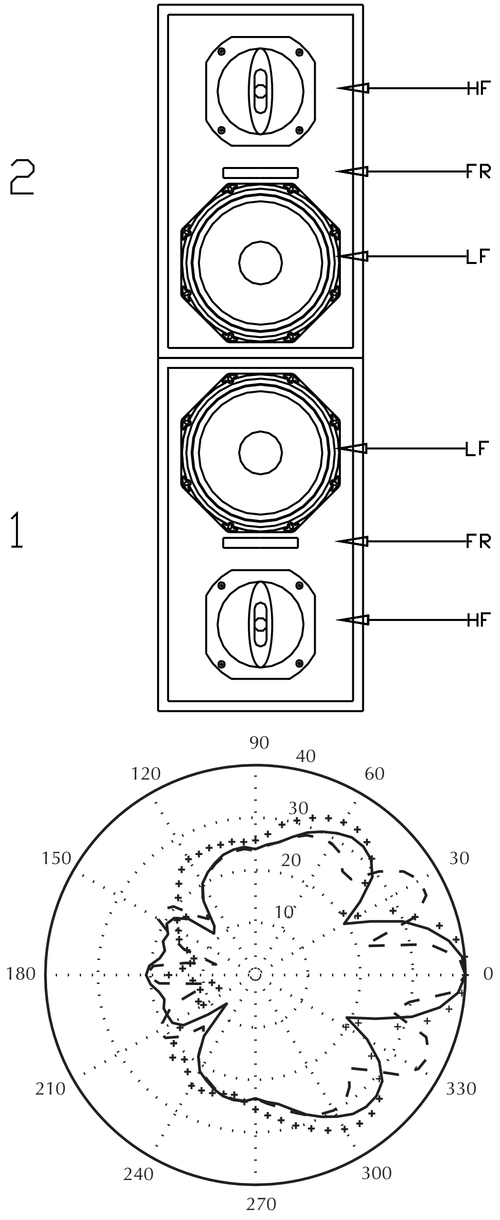
35.2.1.3.3 Configurable Loudspeakers

In the previous sections an overview about the crucial parts of modeling modern loudspeaker systems was given. In turn, the acquisition of complex directivity



Measured data (+), complex data (solid) and magnitude-only data (dashed) at 1 kHz and $^1/_3$ octave bandwidth.

Figure 35-30. Comparison of measurement and prediction for the HF-HF configuration of two two-way loudspeakers.



Measured data (+), complex data (solid) and magnitude-only data (dashed) at 1 kHz and $^1/_3$ octave bandwidth.

Figure 35-31. Comparison of measurement and prediction for the LF-LF configuration of two two-way loudspeakers.

data for individual components creates the basis for another step toward resolving apparent problems on the software side. It allows including system configurability, both electronic and mechanical.

Filter settings of active and passive loudspeakers can now be taken into account in a straightforward way. We can describe the complex radiation function more precisely by including the electronic input $\overline{U}(f)$ into

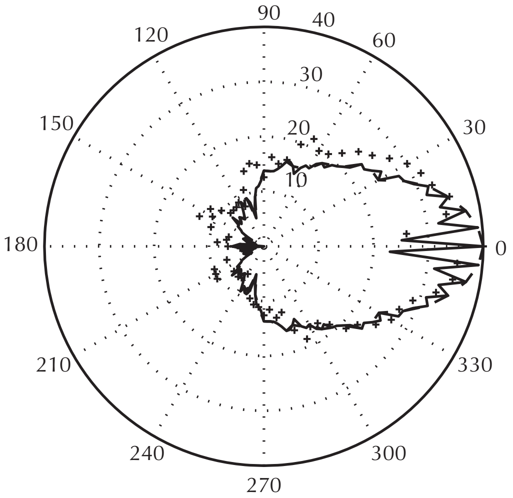


Figure 35-32. Comparison of measurement and prediction for the LF–LF configuration of two two-way loudspeakers, measured data at 5° angular resolution (+), complex data at 5° angular resolution (dashed) and complex data (solid) at 2.5° angular resolution, at 1 kHz and ⅓-octave bandwidth.

the system, the sensitivity of the transducer $\bar{\eta}(f)$, and the filter configuration $\bar{h}(f)$ of the system:

$$\bar{A}(\varphi, \vartheta, f) = \bar{\Gamma}(\varphi, \vartheta, f)\bar{\eta}(f)\bar{h}(f)\bar{U}(f) \tag{35-38}$$

where,

$\Gamma(\varphi, \vartheta, f)$ denotes the angle- and frequency-dependent directivity ratio.

Correspondingly, the coherent pressure sum of several components of a system is expressed by:

$$\bar{p}_{Sum}(\vec{r}, f) = \tag{35-39}$$

$$\sum_n \frac{\bar{\Gamma}_n(\varphi, \vartheta, f)\bar{\eta}_n(f)\bar{h}_n(f)\bar{U}_n(f)}{|\vec{r} - \vec{r}_n|}\exp[-j\vec{k}_n(\vec{r} - \vec{r}_n)]$$

This formulation relates principally to Eq. 35-6 with the equalities of $E_K \sim |\bar{\eta}_n(f)\bar{h}_n(f)|$ and $P \sim |\bar{U}_n(f)|^2$. The loudspeaker properties $\bar{\Gamma}_n(\varphi, \vartheta, f)$ and $\bar{\eta}_n(f)$ will normally be measured and the parameters $\bar{h}_n(f)$ and $\bar{\Gamma}_n(f)$ are defined by the manufacturer or end user. As a result, this concept allows one to model the full response of a multicomponent system under consideration of the given filter settings, may it be a multiway loudspeaker or a digitally steered column. Of course, the effect of changing crossover parameters on the directivity characteristics can be also calculated.[41] An example is shown in Fig. 35-33.

A second step can now be taken as well. The mechanical variability of touring line arrays or clusters can be considered by defining $\vec{r}_n$, either directly by its coordinates or indirectly as a function of user-defined parameters, such as the mounting height of the system and the splay angles between individual cabinets.

35.2.1.3.4 Shadowing and Ray Tracing

It has been pointed out earlier, that for a large-format loudspeaker system, the use of a single point as the origin for ray tracing- or particle-based methods is not adequate. On the other hand, it is not practical to use all individual acoustic sources as origins for the ray-tracing process, given available computing power and the geometrical accuracy of the model. But that is not necessary anyway, since the ray tracing algorithm can be run for subsets or groups of acoustic sources. Therefore representative points have to be found, so-called virtual center points, that can be used as particle sources, Fig. 35-34.

Typical lower-frequency limits for the particle model and the level of detail in common room models suggest ray tracing sources to be spaced apart by about 0.5 to 1 m. In many cases this corresponds to one ray tracing origin per loudspeaker cabinet. While this method of virtual center points is significantly more accurate than using a single source of rays for the whole array, it is still viable with respect to the required computational performance.

35.2.1.3.5 Additional Notes

Some other problems are also automatically resolved by modeling the components of a loudspeaker system separately. For example, the definition of maximum power handling capabilities becomes straightforward. Each component can be described individually by its maximum input level and possibly the frequency response of the test signal. In this respect also the focus of the pro-audio community increasingly shifts from sometimes obscure maximum power values, as defined by the loudspeaker manufacturer, toward the specification of maximum voltage as the entity that is directly measured and applied in modern constant voltage amplifiers.

Finally, one should be aware of the errors made in advanced modeling approaches like the GLL or DLL. It is clear that the acquisition of complex data requires more care and thus engineers will initially see signifi-

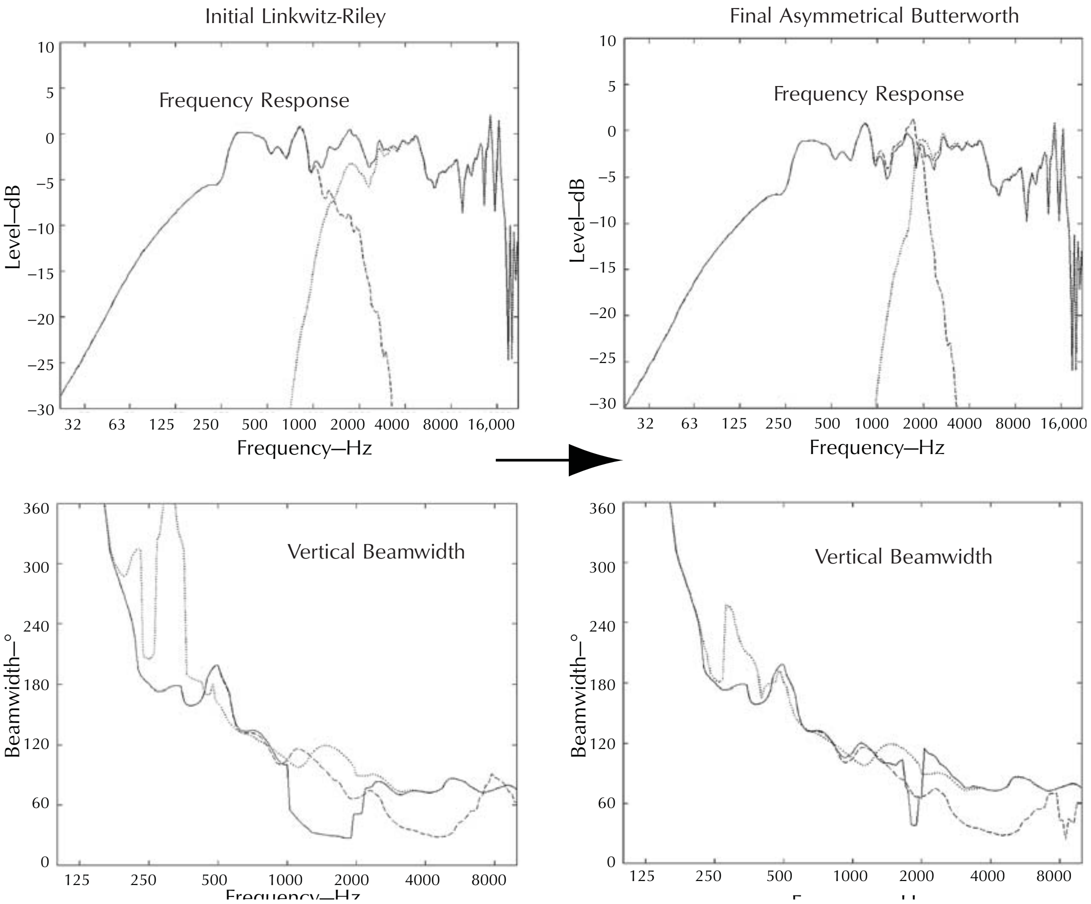


Figure 35-33. Directivity optimization with the prediction software EASE SpeakerLab. Left column shows frequency response and vertical beamwidth of a two-way loudspeaker for initial crossover filter settings, right column shows optimized frequency response and vertical beamwidth. LF unit (--), HF unit (- -), full-range (-).

cant measuring errors, especially with respect to the repeatability of measurements. By refining the measurement setup, using latest measurement technology, and employing data averaging, as well as symmetry assumptions, the data acquisition can usually be improved by an order of magnitude. In addition, it must be emphasized that the variation between samples of the same loudspeaker model may be larger than the measuring error. However, this depends strongly on the manufacturer and its level of quality control.

From the point of view of the simulation software, the best-known practices should be assumed. There is not much sense in limiting the capabilities of an acoustic simulation package because of the quality of the most inexpensive loudspeaker boxes. Like for the geometrical and acoustic model of room, the "garbage in, garbage out" principle holds true for the sound system part of the room as well and the user must be aware of that.

35.2.2 Receiver Simulation

For a complete acoustic model, the acoustic receivers must also be considered. Most important for auralization purposes is to account for the characteristics of the human head and how it influences the sound that reaches the inner ear. Often, simulation software packages also allow utilizing microphone directivity data, in order to be able to image real-world measurement. However, it must be stated that in general the correct implementation of electroacoustic receivers has not nearly received the same level of attention as the sources.

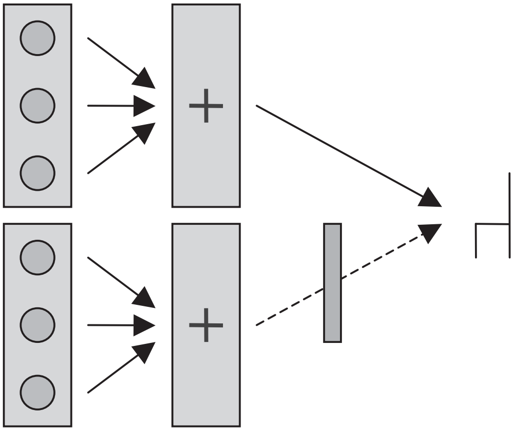


Figure 35-34. Schematic display of shadowing/ray-tracing calculation using virtual center points. Loudspeaker boxes are indicated by gray rectangles, sources indicated by gray circles, virtual centers by black crosses. Only the virtual center points are used for visibility tests.

35.2.2.1 Simulation of the Human Head

Central to incorporating the characteristics of the human head into the simulation results and thus preparing them for final auralization purposes is the head-related transfer function. Typically, this is a data set that consists of two directivity balloons, one for the left ear and a second one for the right ear. Each describes, usually by means of complex data, how the human head and the outer part of the ear change the incoming sound waves as they arrive at the ear. It is critical for a satisfactory binaural auralization that the signal for each ear is weighted with an appropriate angle- and frequency-dependent directivity function.

The acquisition of measurement data for the human head is not a trivial matter. Since real human heads cannot be measured directly, a so-called dummy head has to be built or in-ear microphones have to be used, see Section 35.1.5.1 Human Ears. Each ear of a dummy or a real head is equipped with a microphone. Balloon measurements are made similar to loudspeaker balloon measurements, only that the locations of source and receiver are reversed and a stereo set of data files is obtained.[27]

Recent research[42] has shown that the inclusion of the human torso into the HRTF also has significant effect on the quality of the binaural reproduction. Even more so, auralization results of highest quality can be obtained utilizing a head-tracking system and a set of HRTF balloons, where each pair of balloons describes the transfer function for the left and right ear for a particular angular position of the human head relative to the human body. This data can then be employed to auralize impulse responses of either a measured or simulated environment with speech and music contents.

35.2.2.2 Simulation of Microphones

The need for inclusion of microphones in acoustic simulation software has several reasons. On the one hand, to be able to compare measurements with computational results, the frequency response and the directivity characteristics of the microphone have to be taken into account. On the other hand, the possibility to simulate either recording or reinforcement of a talker or musician is of practical interest too. For example, by varying the location and orientation of the pick-up microphones the coverage can be optimized. Finally, by including microphones, it becomes possible to simulate the entire chain of sound reinforcement, from the source over the microphone to the loudspeaker and back to the microphone. Only this enables the prediction of feedback and to estimate the potential gain before feedback.

However, the acquisition and distribution of microphone data must still be considered in its infancy. Available data consists largely of octave-based magnitude-only data that assumes axial symmetry. Measurement techniques vary significantly among microphone manufacturers and measuring conditions, such as the measurement distance, are not standardized and often not even documented. Therefore most users of simulation programs do not consider implementing microphone data into their models, or if so, they use generic data based on ideal directional behavior, like cardioid or omnidirectional patterns.

There are several more issues that inhibit the widespread acquisition, acceptance, and use of microphone data.

- First, especially the measurement distance is important with respect to the acquisition of the data and its application in the software domain. A lot of microphones exhibit the so-called proximity effect, that is, the property that their frequency response and directivity function change depending on the shape of the incident wave front. This effect is most visible if the acoustic source is within a few meters, range of the microphone and thus the wave front cannot be considered as plane anymore.
- Secondly, we described earlier with respect to loudspeaker data, that it is important to preserve configurability also in the software domain. In this regard, switchable multipattern microphones have to be

taken into account when developing a fully descriptive data model.

- The use of combined microphones is also wide spread. In particular, multichannel receivers, such as dummy heads, coincidence recording microphones, or B-format receivers, need to find an appropriate representation in the simulation software.
- Another issue of concern is the acquisition of phase data. The impact of neglecting the phase of the loudspeaker on the simulation of its performance is well known. But not much research has happened in that respect regarding microphones. Nevertheless, it is clear that under special circumstances like in feedback situations or for the electronic combination of microphone signals (e.g., two active microphones on lecterns) phase plays an important role.
- Finally, of course, it must be stated that the usability of microphone data has its limitations depending on the application of the particular model. Compared to installation microphones typical handheld microphones have different properties. The data that is needed and that can be acquired may differ accordingly.

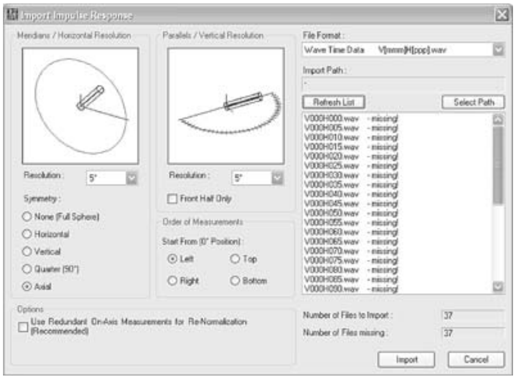
Recently an advanced data model was proposed that is able to resolve many of the issues listed above.[43] Basically, it suggests using a similar approach like the loudspeaker description language (GLL) introduced earlier, namely to describe receiver systems in a generalized, object-oriented way. This means especially that:

- Microphone data files should at least include far-field data (plane wave assumption), but can also contain proximity data for various near-field distances.
- A microphone model can consist of multiple receivers, that is, acoustic inputs, and can have multiple channels (electronic outputs).
- A switchable microphone should be represented by a set of corresponding data subsets.
- Impulse response or complex frequency response data should be utilized to describe the sensitivity and the directional properties of the microphone as appropriate.

Fig. 35-35 is an example for an import function in the new EASE Microphone Database software.

35.3 Tools of Simulation

Today, an acoustic CAD program must be able to predict all needed acoustic measures exactly enough. A 100% forecast is certainly impossible but the results of a computer simulation must come close to reality (errors generally equal or less than 30%). Then it becomes possible that the acoustic behavior of a facility can be made audible by so-called auralization. (One will listen to sound events just performed by means of the computer.) The following will give a short introduction of the possibilities of computer simulation today.



File Format : Wave Time Data V[mmm]H[ppp].wav

Import Path : -

Resolution : 5°

Resolution : 5°

○ None (Full Sphere)
○ Horizontal
○ Vertical
○ Quarter (90°)
● Axial

☐ Front Half Only

● Left ○ Top
○ Right ○ Bottom

☐ Use Redundant On-Axis Measurements for Re-Normalization (Recommended)

V000H000.wav - missing
V000H005.wav - missing
V000H010.wav - missing
V000H015.wav - missing
V000H020.wav - missing
V000H025.wav - missing
V000H030.wav - missing
V000H035.wav - missing
V000H040.wav - missing
V000H045.wav - missing
V000H050.wav - missing
V000H055.wav - missing
V000H060.wav - missing
V000H065.wav - missing
V000H070.wav - missing
V000H075.wav - missing
V000H080.wav - missing
V000H085.wav - missing
V000H090.wav - missing

Number of Files to Import : 37

Number of Files missing : 37

Figure 35-35. Import routine in EASE Microphone Database Software.

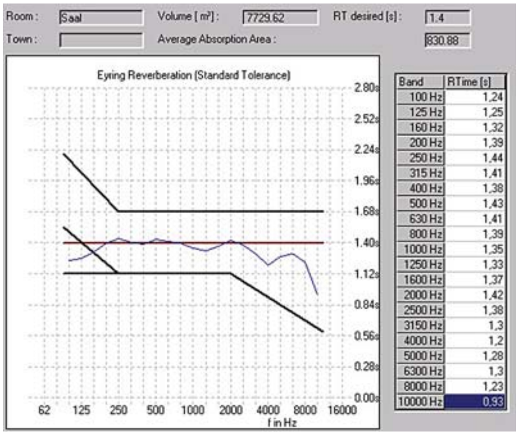
35.3.1 Room Acoustic Simulation

35.3.1.1 Statistical Approach

Based on simple room data and the associated surface absorption coefficients, a computer program is able to calculate the reverberation time according the Sabine and Norris-Eyring equations, see Section 7.2.1.1. On the other side measured values must be usable directly in such a program. Calculation of the early decay time (EDT) should be possible too.

A comprehensive database of country-specific and international wall materials and their absorption coefficients is part of the program. This database should be accessible to allow the user to import and enter data from other textbook sources or measurements. Because most of the needed scattering coefficients are not available in textbooks a computer program should allow deriving values even by rules of thumb.

A set of frequency-dependent target reverberation times should be available for entering into the simulation program so that the room models, calculated (or real-world measured) RT_{60} times can be compared with the target values. The program should then indicate (for each selected frequency band) the calculated (or measured) Fig. 35-35 time versus the target RT_{60} time and list the number of excess or deficient RT_{60} times for each band relative to the target values within a range of tolerance, Fig. 35-36.



Room: Saal
Volume [m³]: 7729.62
RT desired [s]: 1.4
Town:
Average Absorption Area: 830.88

Band	RTime [s]
100 Hz	1,24
125 Hz	1,25
160 Hz	1,32
200 Hz	1,39
250 Hz	1,44
315 Hz	1,41
400 Hz	1,38
500 Hz	1,43
630 Hz	1,41
800 Hz	1,39
1000 Hz	1,35
1250 Hz	1,33
1600 Hz	1,37
2000 Hz	1,42
2500 Hz	1,38
3150 Hz	1,3
4000 Hz	1,2
5000 Hz	1,28
6300 Hz	1,3
8000 Hz	1,23
10000 Hz	0,93

Figure 35-36. RT_{60} chart with tolerance range.

The graph of RT_{60} times should allow plotting multiple RT_{60} values within a single graph, so as to show the impact of various audience sizes, proposed and/or alternative room treatments, etc., on the RT_{60} time. An option must allow plotting a grayed or dashed area as the desirable range of reverberation times for a particular project, against which the measured or calculated RT_{60} values can be referenced.

35.3.1.2 Objective Room-Acoustic Measures

The simplest way to obtain objective measures is to use the direct sound of one or more sources and calculate the reverberation level of the room by means of the reverberation time equations assuming the room follows a statistically even distributed sound decay (homogeneous, isotropic diffuse sound field, that is, the RT_{60} is constant over the room). From these calculations it is possible to derive the direct sound and the diffuse-sound levels and consequently a range of objective acoustic parameters, see Section 7.1. It goes without saying that this requires the acoustical conditions of the room to show a statistically regular behavior (frequency response of the reverberation time that is independent of the location considered in the room). In practice, however, such behavior will hardly be found. For this reason one tends to qualify such data as having only a preliminary guideline character and to have it confirmed by additional detailed investigations.

35.3.2 Ray Tracing or Image Modeling Approach

35.3.2.1 Preliminary Remark

There are several ways to calculate the impulse response of a radiated sound event. The widest-known method is the image-source algorithm. Worth mentioning at this point are also the ray-trace method, which was first known in optics, and other special procedures like cone tracing or pyramid tracing. Nowadays these procedures are more often than not used in a combined form as so-called hybrid procedures.

35.3.2.2 Image Modeling

With image modeling there is a source and a receiving point selected. Then a deterministic search of all image sound sources of different orders is started to compute the impulse response, Fig. 35-37.

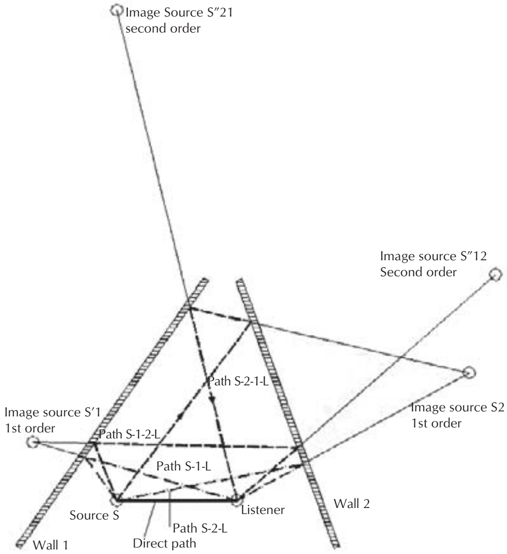


Figure 35-37. Ray calculation with image model algorithm.

In the image modeling method a receiving point is used instead of a counting balloon (in contrast to classical ray tracing). Frequency response and interference effects (including phase investigations) are also easily calculated.

This method is very time consuming and the calculation time is proportional to N^i with: N = number of model walls and i = the order of wall bounces.

So one gets usable results for models with $N < 50$ and $i < +6$. For larger models and more complicated investigations the next method is more advantageous.

35.3.2.3 Ray Tracing

In contrast to image modeling, here the path of a single sound particle radiated under a random angle into the room along a ray is followed. All surfaces are checked to find the reflection points (with or without absorption or diffusion). The tracing of the single ray is terminated when the remaining sound energy has decreased to a certain level or when the particle hits an appropriately arranged counting balloon with a finite diameter, typically at the location of a listener in the room, Fig. 35-38.

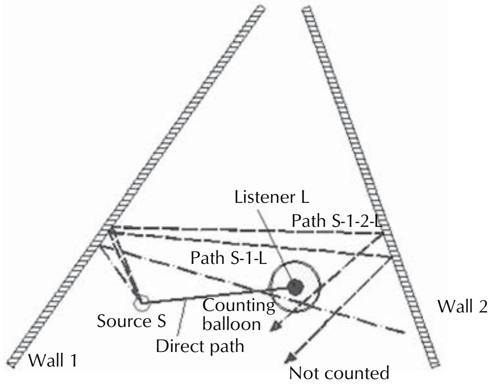


Figure 35-38. Ray calculation with ray-tracing algorithm.

Phase considerations are not possible directly, but can be derived if an image model routine is run that retraces the last ray after intercepting the counting balloon or if the ray retains the information about its sequence of reflection points throughout the process.

This method runs significantly faster and the calculation time is only proportional to the number N of the model walls. Ray-tracing methods can be even faster, if they are based on logarithmic search for the intersection points (~ log$N$).

35.3.2.4 Cone Tracing

This method is used in various CAD programs. Its advantage is the directed ray radiation over the different room angles, Fig. 35-39.

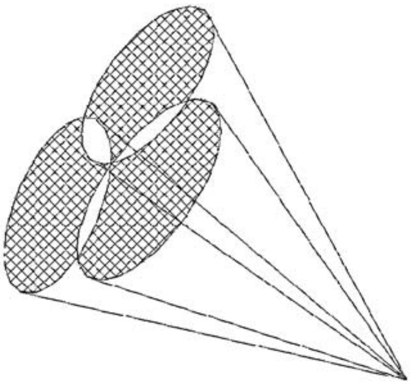


Figure 35-39. Ray radiation in cones.

Because of these cones, fast ray calculations can proceed. The fact that the cones do not cover the source "sphere" surface completely turns out to be a disadvantage. It is necessary to overlap adjacent cones and an algorithm is required to avoid multiple detections or to "weight" the energy so that the multiple contributions produce (on average) the correct sound level. Some famous conical beam tracers are known, implementing different techniques to correct this point.[44,45,46]

35.3.2.5 Pyramid Tracing

This method was introduced by Farina in the program "Ramsete" in 1995.[47]

Farina demonstrated that the pyramid beams do not suffer from the cone-trace overlap, as adjacent pyramids cover perfectly the source sphere, Fig. 35-40.

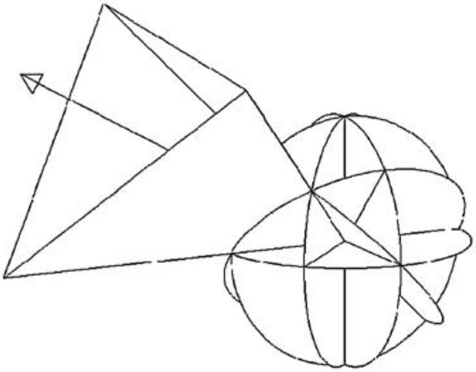


Figure 35-40. Ray radiation in pyramids.

Originally a subdivision of the surface in triangles was made by subsequent subdivisions of the 8 octants of the sphere: according to Farina "this way the number of pyramids generated can be any power of 2, and all of them have almost the same base area, giving a nearly isotropic sound source."

35.3.2.6 Room-Acoustic Analysis Module-(AURA)

To illustrate these methods, as an example the new hybrid ray-tracing algorithm AURA will be explained in more detail in the following.

Based on CAESAR,[48] the AURA algorithm[49] calculates the transfer function of a room for a given receiver point using the active sound sources. For this purpose a hybrid model is employed that uses an exact image source model for early specular reflections and an energy-based ray-tracing model for late and scattered reflections. The transition between the two models is determined by a fixed reflection order.

The ray-tracing model utilizes a probabilistic particle approach and can therefore be understood as a Monte-Carlo model. At first, the sound source emits a particle in a randomly selected direction with a given energy. The particle is then traced through the room until it either hits a boundary or a receiver or its time of flight reaches the user-defined cut-off time. When the particle hits a boundary it is attenuated according to the surface material and its direction is adjusted according to the reflection law. An essential assumption of this Monte-Carlo approach is that attenuation due to air or surface reflections is taken into account as a reduction of particle energy, while the propagation loss over distance is indirectly covered by the reduced detection probability for individual particles with increasing distance and fixed receiver sizes.

Per receiver and simulated frequency, a so-called echogram is created that contains energy bins linearly spaced in time. When a receiver is hit, the energy of the detected particle is added to the bin that corresponds to the time of flight. Also, as a separate step, the contributions from the image source model are included. The particle model accounts for scattering in a probabilistic way. Whenever a particle hits a surface, the material absorption part is subtracted from its energy. Then, a random number is generated and depending on the scattering factor, the particle is either reflected geometrically or it is scattered under a random angle based on a Lambert distribution. After that the particle is traced until it hits a receiver or a wall again.

For room acoustic models brute-force ray tracing, that is, testing all model walls or wall triangles for intersection, is often impractical since computation time scales linearly with the number of triangles. Improved performance is obtained by structuring triangle data such that each ray is tested for intersection only with a subset of triangles. Current methods are based on two main strategies: hierarchical bounding volumes (HBV)[50] and space partitioning.[51] In the former case, a hierarchy of simple bounding volumes (such as spheres) is constructed, where a particular volume may include either a number of smaller child-volumes or actual triangles. A ray is tested for intersection starting at the top of the hierarchy, such that a particular child-volume is only tested if the parent was hit. The cost of ray-bounding volume intersection is small, and the resulting computation scaling with the number of triangles is approximately logarithmic. In space partitioning schemes, the physical space where the triangles reside is partitioned into smaller cells or so-called voxels. Rays are followed through adjacent voxels and tested only against triangles pertaining to those voxels. The partitioning may be uniform or more complex—e.g., hierarchical, adaptive, etc.

Previous studies indicate that no particular ray-tracing acceleration structure is obviously the most efficient, since the total computation cost depends both on algorithm and hardware implementation.[52] Whereas highly refined hierarchical acceleration schemes may require less intersection tests, the associated data structures are nonuniform (i.e., hard to parallelize), involve traversal of nonlocal data structures, and as such are less suitable for cache and vector processing optimizations as available on modern processors and graphics cards. On the other hand, space partitioning methods, in particular those involving simple data structures like uniform grids, are more suitable to efficient implementation on vector processing elements.

In AURA a uniform grid ray-tracing algorithm is implemented similar to Amanatides and Woo.[53] A 3D uniform grid is assigned to the simulation box and each triangle is associated with every cell having a common interior point with it. The grid spacing in every direction is determined automatically via an empirical formula: the number of cells on each axis is proportional to the square root of the total number of triangles, and to the box length along the axis divided by the average box dimension (since in general the triangles form approximately a 2D shell, such a formula matches the average cell dimension to the average triangle dimension). Up to sixty four cells per axis are allowed, in order to limit memory requirements. Given a ray specified by an origin and direction vector, a fast grid traversal algorithm computes the next grid cell intersected by the ray. Each triangle associated with this grid cell is then tested for intersection with the ray. No particular optimization is done to avoid duplicate ray-triangle tests when one triangle spans multiple voxels. Thus a ray-triangle intersection is only considered if it occurs within the boundaries of the current cell. The grid traversal continues until a hit point is found or the ray exits the simulation box.

The software implementation in AURA was carefully designed to facilitate vector optimization on SIMD-capable processors—i.e., to minimize branching and optimize instruction scheduling—in particular it is easily transferable to programmable graphics processing units. Using the Intel C++ optimizing compiler on a Pentium4 processor, the grid algorithm applied to realistic test cases with more than 10,000 triangles can be up to five times faster than a typical HBV method and, of course, orders of magnitude faster than the linear search method.[54]

35.3.2.7 Features of All These Methods

All of the ray-tracing or image modeling methods that calculate impulse responses have to take into account the directivity of the sound sources and the absorptive and scattering characteristics of the surfaces encountered en route from the source to the receiving point.

The design program must allow the user to designate specific surfaces/planes as being reflective or nonreflective. This will make it possible to simulate not only sound-reflecting walls, but also simplified floor planes—i.e., which in reality are complex shapes such as seating areas or orchestra stages or pits. At present these methods use statistical absorption factors that are readily available instead of angle-dependent ones (for the latter no sources are available in textbooks), as well as some diffusion factors estimated by rule of thumb and/or specially measured diffusion factors. The diffraction behavior is still in the academic stage and some program approaches are using FEM or BEM methods[26,35], see Section 35.1.4.2. and Fig. 7-46. Additionally the dissipation of sound energy in air—i.e., the frequency-dependent air attenuation—must be considered too.

A library of potential natural sound sources must be available, such as the human voice and various orchestral instruments/sections to go along with the electroacoustic sources/loudspeakers that should include the sound power level and directivity of these sources/loudspeakers. As a result of all these calculations you get impulse responses or energy-time curves as shown in the following figures.

The program CATT-Acoustic[36] shows the complete echogram with all input data (room, loudspeaker, listener position, frequency) and presents all resulting room acoustic measures this way, Fig. 35-41. With EASE and AURA it looks different, Fig. 35-42.

The calculated energy-time curve should be able to be stepped through reflection by reflection, with the appropriate rays and surfaces being highlighted to indicate the ray's path and the surfaces it encounters en route from source to receiver, Fig. 35.43.The software should indicate median/lateral/horizontal positioning of energy arrivals (and relative magnitude as well) at the receiver's location, Fig. 35-44.

Additionally, a simulation program should provide the capability to calculate early/late energy ratios. It is important to be able to set the early/late transition time and also to select the cutoff time for the late energy integral, Fig. 35-45.

The software's ray-tracing or image modeling method of deriving an energy-time curve should provide the ability to indicate interaural cross-correlation (IACC) as well as lateral energy coefficient predictions at specified listener positions.

35.3.3 Auralization

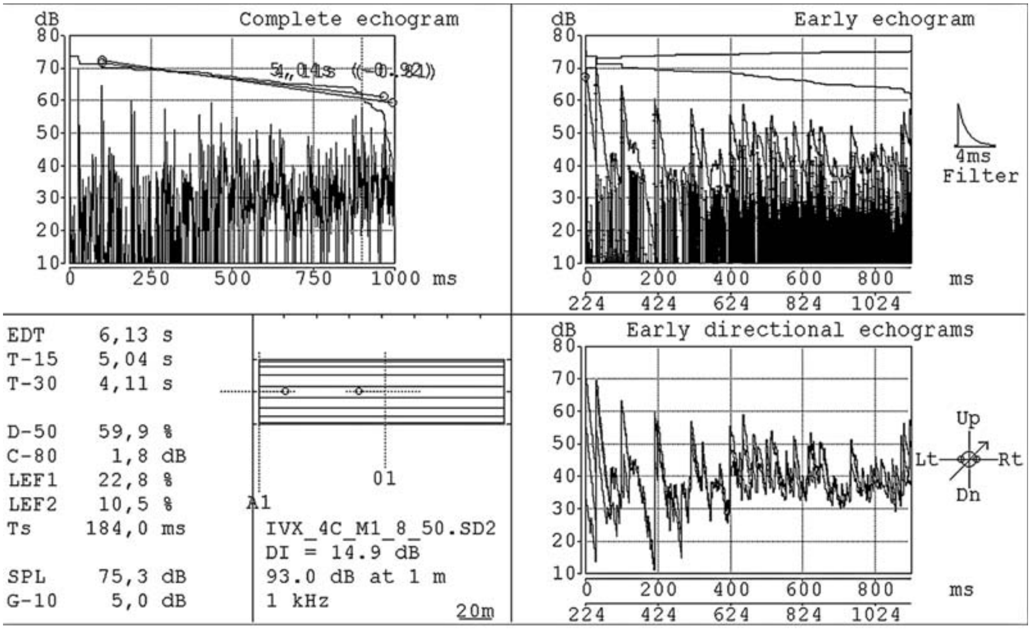
The simulation program must have the ability to transfer the calculated impulse response curve to a postprocessing routine that will be used to auralize the room time/energy data with anechoic music or speech source material. Of course the routine must generate a binaural data file in WAV-format or other computer sound file format in common use, Fig. 35-46.

35.3.4 Sound Design

35.3.4.1 Aiming

Aiming the individual loudspeakers is an important operation insuring the proper spatial arrangement and orientation of the sound reinforcement systems. Once the corresponding room or open-air model is at hand and the mechanical and acoustical data of the loudspeaker systems is exactly known, these systems are approximately positioned and then one may begin with the fine tuning of the same. A modern simulation program uses a kind of isobeam/isobar method to initially aim the loudspeakers, preferably utilizing the –3 dB, –6 dB or –9 dB contours.

Fig. 35-47 shows various types of projection of the –3, –6, and –9 dB curves into the room. On audience areas one can then also see superposed aiming curves for multiple loudspeakers, Fig. 35.48.



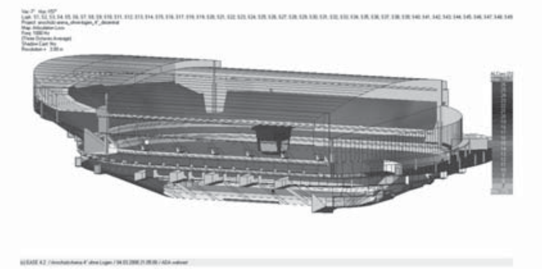
EDT 6,13 s
T-15 5,04 s
T-30 4,11 s

D-50 59,9 %
C-80 1,8 dB
LEF1 22,8 %
LEF2 10,5 %
Ts 184,0 ms

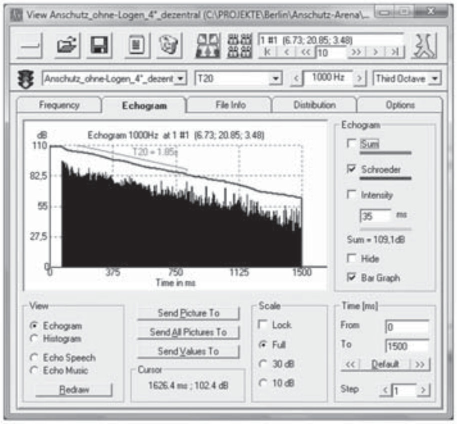
SPL 75,3 dB
G-10 5,0 dB

IVX_4C_M1_8_50.SD2
DI = 14.9 dB
93.0 dB at 1 m
1 kHz

Figure 35-41. Echo and data plot in CATT acoustics.



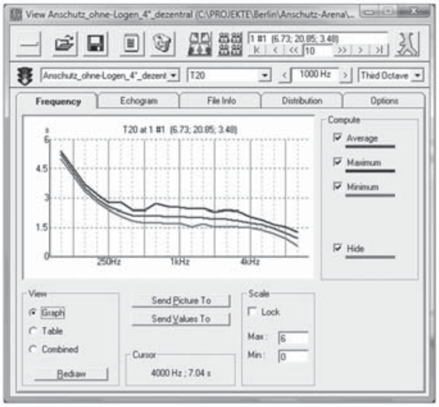
A. 3D mapping inside a model.



View Anschutz_ohne-Logen_4"_dezentral (C:\PROJEKTE\Berlin\Anschutz-Arena\...
1 #1 (6.73; 20.85; 3.48)
Anschutz_ohne-Logen_4"_dezent
T20
1000 Hz
Third Octave
Echogram 1000Hz at 1 #1 (6.73; 20.85; 3.48)
T20 = 1.85s
☐ Sum
☑ Schroeder
☐ Intensity
35 ms
Sum = 109,1dB
☐ Hide
☑ Bar Graph
View
● Echogram
○ Histogram
○ Echo Speech
○ Echo Music
Cursor
1626.4 ms ; 102.4 dB
Scale
☐ Lock
● Full
○ 30 dB
○ 10 dB
Time [ms]
From 0
To 1500
Step 1

10
Frequency
Echogram
File Info
Distribution
Options
dB
110
82,5
55
27,5
0
375
750
1125
1500
Time in ms
Send Picture To
Send All Pictures To
Send Values To
Redraw
Default

B. Echogram in EASE-AURA.



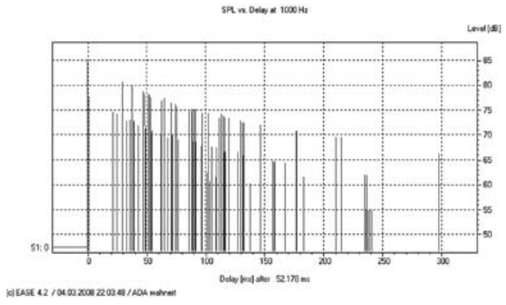
View Anschutz_ohne-Logen_4"_dezentral (C:\PROJEKTE\Berlin\Anschutz-Arena\...
1 #1 (6.73; 20.85; 3.48)
Anschutz_ohne-Logen_4"_dezent
T20
1000 Hz
Third Octave
T20 at 1 #1 (6.73; 20.85; 3.48)
Compute
☑ Average
☑ Maximum
☑ Minimum
☑ Hide
View
● Graph
○ Table
○ Combined
Cursor
4000 Hz ; 7.04 s
Scale
☐ Lock
Max: 6
Min: 0

10
Frequency
Echogram
File Info
Distribution
Options
s
6
4.5
3
1.5
0
250Hz
1kHz
4kHz
Send Picture To
Send Values To
Redraw

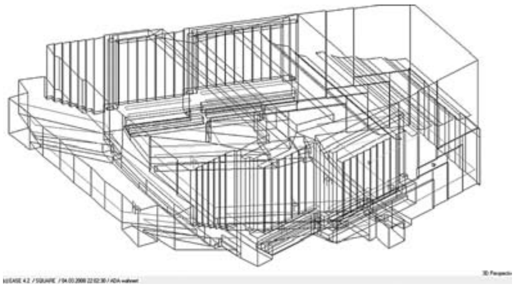
C. Reverberation Time plot in EASE-AURA.

D. 2D mapping.

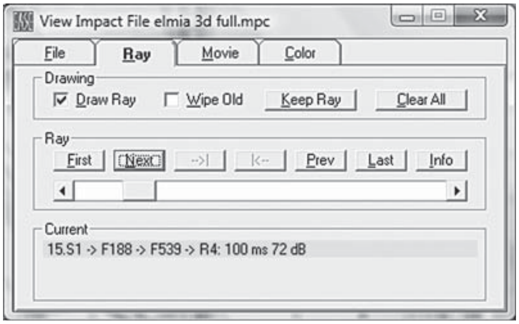
Figure 35-42. Echo and data plot in EASE 4.2



A. Reflectogram.



B. Ray visualization in an EASE model.



☑ Draw Ray ☐ Wipe Old

15.S1 -> F188 -> F539 -> R4: 100 ms 72 dB

C. Control folder for visualization.

Figure 35-43. Reflectogram and ray visualization in EASE 4.2.

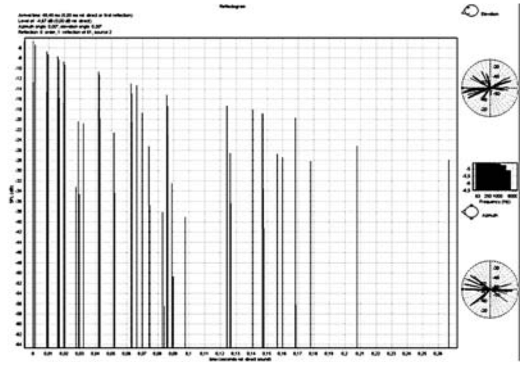


Figure 35-44. Reflectogram and Impact hedgehogs in ODEON 9.0.

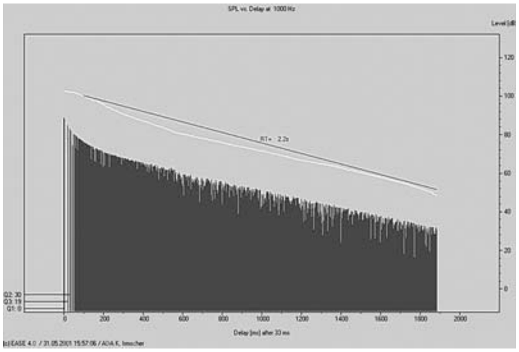


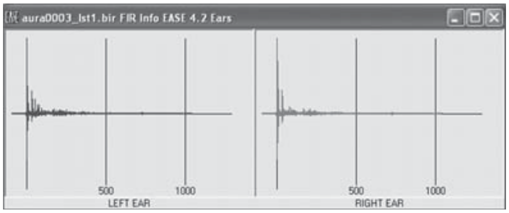
Figure 35-45. Reflectogram with tail and Schroeder plot.

35.3.4.2 SPL Calculations

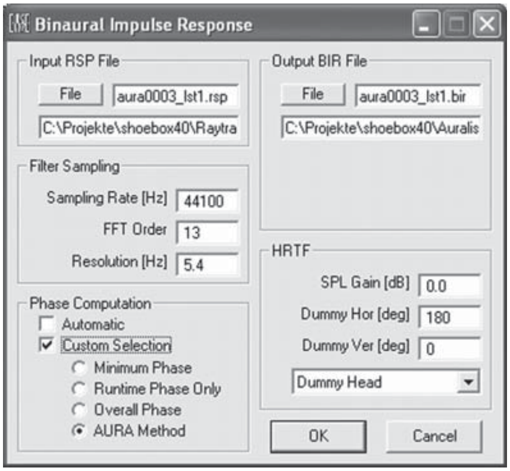
After the loudspeakers have been correctly aimed, one may begin to calculate the sound-level conditions attainable by these. The first results are given for the direct sound pressure level (SPL). As long as we predict a good direct sound coverage over the listener area we have also to expect perfect intelligibility numbers, of course under the condition that the reverberation level is not too high.

A complex summation—phase conditions including travel-time differences should be included—has to be used as the standard method of calculating the direct SPL. This method is exact for a planar wave, but only an approximation for the superposition of waves with different propagation directions. But the complex sound pressure components of different coherent sources must first be added and afterward squared to obtain SPL numbers. In so-called DLL or GLL approaches one always calculates the complex sum of all sources in the array.

Today simulation programs are usually still only analyzing programs, capable of calculating which levels can be obtained by which loudspeakers and under which acoustical conditions. But questions are more and more asked the other way around. The program of the future also should query the user for a desired average SPL of the system, and automatically adjust the power provided to each loudspeaker—with a warning when the power required exceeds the capabilities for the loudspeaker, based on the desired SPL of the design, the sensitivity and directivity of the loudspeaker, the distance of throw, and the number of loudspeakers. This presupposes, of course, new algorithms that in most of the simulation programs are just being developed.



A. Calculated binaural impulse responses.



Binaural Impulse Response

Input RSP File

File aura0003_lst1.rsp

C:\Projekte\shoebox40\Raytra

Output BIR File

File aura0003_lst1.bir

C:\Projekte\shoebox40\Auralis

Filter Sampling

Sampling Rate [Hz] 44100

FFT Order 13

Resolution [Hz] 5.4

Phase Computation

☐ Automatic

☑ Custom Selection

○ Minimum Phase

○ Runtime Phase Only

○ Overall Phase

● AURA Method

HRTF

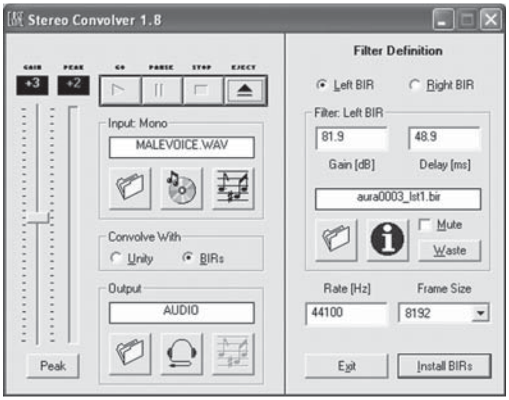
SPL Gain [dB] 0.0

Dummy Hor [deg] 180

Dummy Ver [deg] 0

Dummy Head

B. Convolution with HRTF files.



Stereo Convolver 1.8

GAIN +3

PEAK +2

Input: Mono

MALEVOICE.WAV

Convolve With

○ Unity ● BIRs

Output

AUDIO

Filter Definition

● Left BIR ○ Right BIR

Filter: Left BIR

Gain [dB] 81.9

Delay [ms] 48.9

aura0003_lst1.bir

☐ Mute

Rate [Hz] 44100

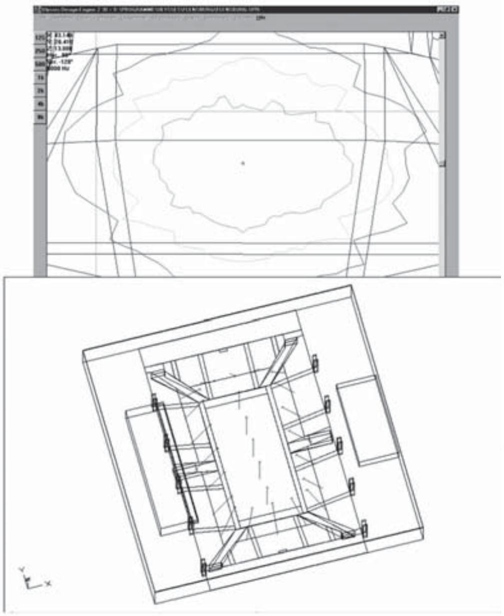
Frame Size 8192

C. Real time stereo convolver unit in EASE 4.2.

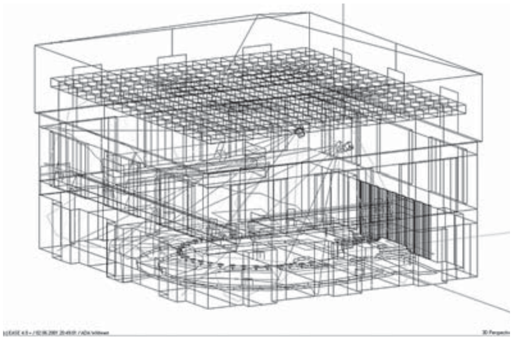
Figure 35-46. Auralization.

In Fig. 35-49 the level-time-frequency-behavior of a loudspeaker cluster at a chosen listener seat in a room is shown by a simulated waterfall diagram.

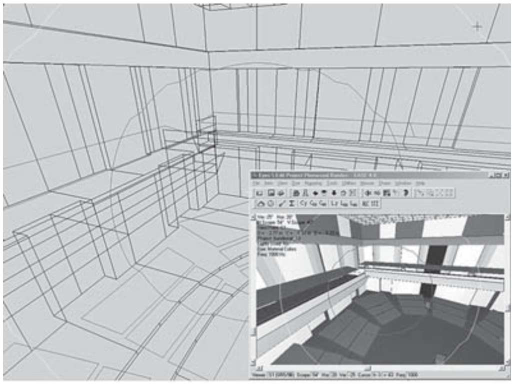
The target of all the efforts is to cover the whole audience area(s) evenly with musically pleasing and intelligible sound, while providing sound pressure levels suitable for the intended purpose.



A. 3D presentation in ULYSSES 2.3



B. 3D presentation in EASE 4.0 wireframe model.



C. 3D presentation in EASE 4.0 rendered model.

Figure 35-47. 3D aiming presentations in simulation programs.

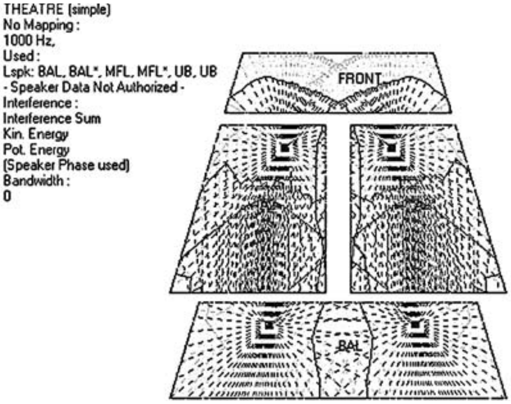


Figure 35-48. 2D aiming mapping in EASE 4.2.

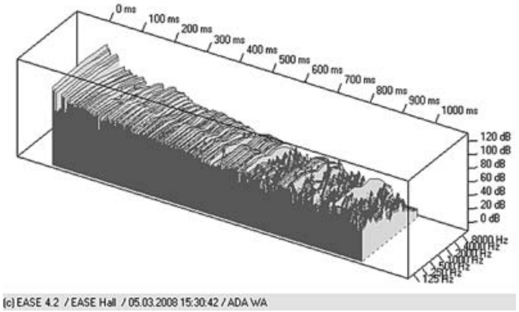


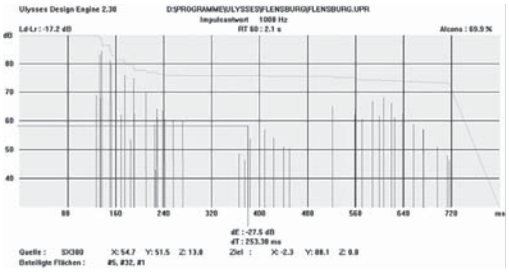
Figure 35-49. Waterfall presentation in EASE 4.2.

All simulation programs today are widely lacking the algorithms required for computing the acoustical feedback. These, however, should soon become available, since the available microphone data is similar in structure to that of loudspeakers, see Section 35.2.1.2. Then it will be possible to calculate the maximum and nominal acoustical gain based on the parameters of microphone/loudspeaker/listener/talker, which take into account the number of microphones and/or the arrangement of loudspeakers.

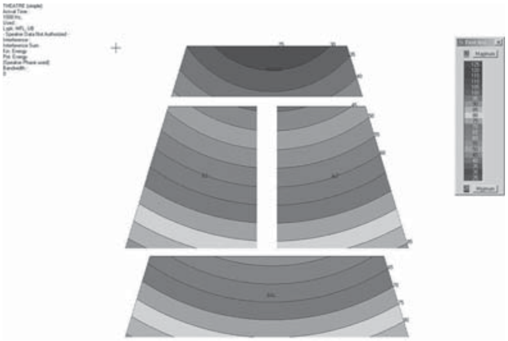
35.3.4.3 Time Arrivals, Alignment

A graph of time arrivals (direct, direct + reflected, reflected only) should allow the user to show the first energy arrival as required by the design, to adjust a signal delay loudspeaker to bring the loudspeakers into synchronicity, and to realize an acoustic localization of an amplified source—via distance and the HAAS effect, see Figs. 35-50A and B.

Matters are often complicated by special requirements such as localization, stereo imaging, etc. Simulation programs allow determining the first wave front as well as calculating initial time delay gaps or echo detections (c.f. in this respect Figs. 35-51A to C).



A. Reflectogram in ULYSSES 2.3.



B. Delay pattern of first arrival in EASE 4.2.

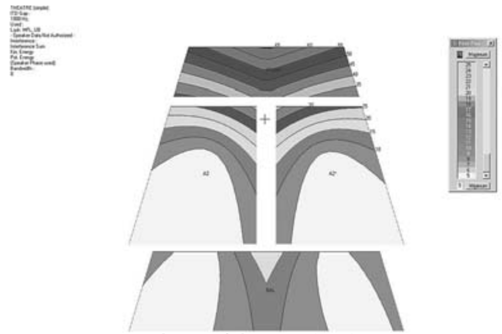
Figure 35-50. Delay presentations in simulation programs.

Predicted array lobing patterns of arrayable loudspeakers should be displayed by simulation programs, with the ability to provide signal delay and/or move the appropriate loudspeakers to attempt to bring the array into acoustic alignment. A program today will have the ability to provide signal delay to the individual loudspeakers to align them in time. The corresponding sound pressure calculations will take into account either measured phase data for the individual loudspeakers or the run-time phase if phase differences among the components can be neglected.[56]

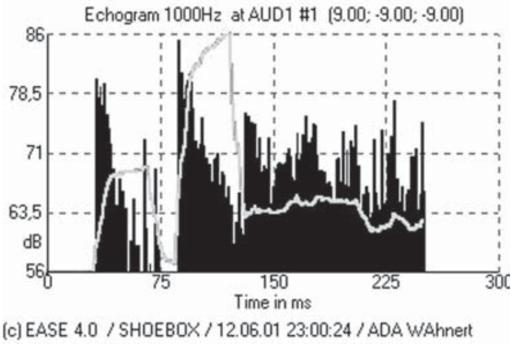
Fig. 35-52A and B shows the frequency response of nonaligned and aligned loudspeaker groups, simulated by EASE/ULYSSES.

35.3.4.4 Mapping, Single-Point Investigations

Once the aiming, power setting, and alignments are completed, the program should provide a colored visual coverage map of the predicted sound system performance. This coverage map must take into account the properties of the loudspeakers as well as the impact of reflecting or shadowing planes, and provide the following displays at a minimum:

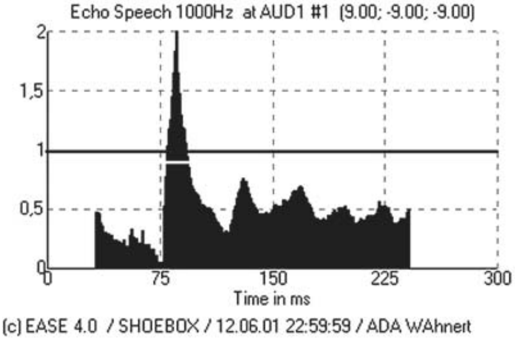


A. Initial time delay gap (ITD) mapping.



(c) EASE 4.0 / SHOEBOX / 12.06.01 23:00:24 / ADA WAhnert

B. Echogram in weighted integration mode.

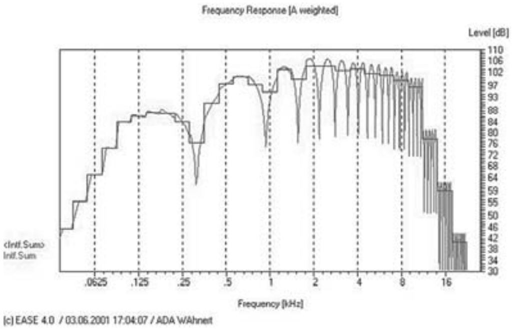


(c) EASE 4.0 / SHOEBOX / 12.06.01 22:59:59 / ADA WAhnert

C. Echo detection curve for speech.

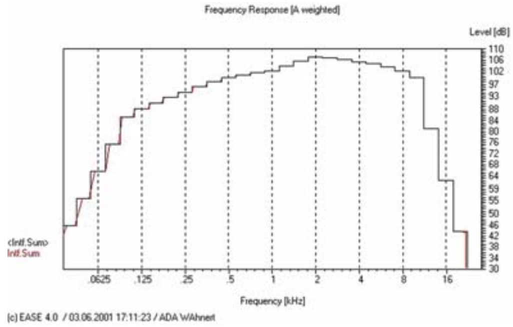
Figure 35-51. Echo detection in EASE 4.2.

- Predicted sound pressure level, viewed at octave or ⅓ octave band frequencies, and at an average of these frequencies, Figs. 35-53A to C.
- Predicted intelligibility values (in the 2 kHz octave band, or the weighted average of 500 Hz to 4 kHz octave band data), listed in STI or RASTI values, Figs. 35-54A and B.
- Predicted acoustic measures (for octave or ⅓ octave band frequencies), listed in C80, C50, %*Alcons*, center time, strength, or other values according to ISO Standard 3382 (compare Fig. 35-55A and B).



(c) EASE 4.0 / 03.06.2001 17:04:07 / ADA WAhnert

A. Non-aligned cluster.



(c) EASE 4.0 / 03.06.2001 17:11:23 / ADA WAhnert

B. Aligned cluster.

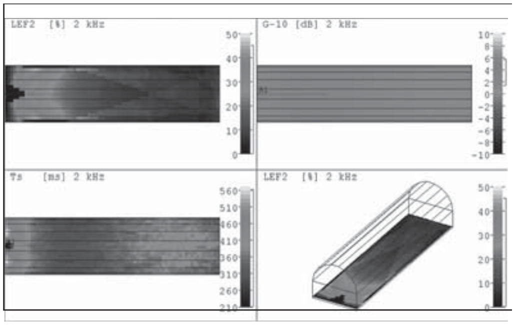
Figure 35-52. Frequency response of loudspeaker cluster.

35.4 Verification of the Simulation Results

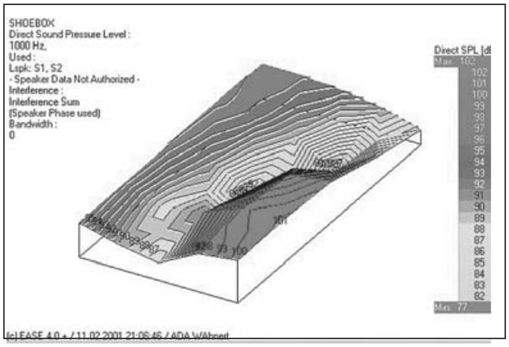
After the simulation, the practical design, and the installation, it is important to check the results and to compare them with the prediction. For this purpose tools we developed during the last 20 years:

- The most famous TEF 10, 12, and 20 by Crown (later Gold Line).
- MLSSA by DRA Laboratories.
- SMAART by SIA Soft.
- WinMLS by Morset Sound Development.
- DIRAC by Brüel & Kjær.
- SpectraLAB by Sound Technology Inc.
- EASERA by AFMG Berlin.
- EASERA SysTune by AFMG Berlin.

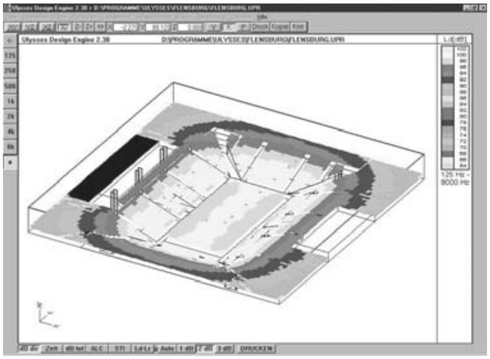
All measurements with predefined excitation signals generally utilize two or more ports. The input port of the system under test (DUT) is fed with an excitation signal, generated by the analyzer. Fig. 35-56 shows the block-diagram for a modern software-based four-port measurement tool including the needed AD/DA converter.


LEF2 [%] 2 kHz
G-10 [dB] 2 kHz
Ts [ms] 2 kHz
LEF2 [%] 2 kHz
50
40
30
20
10
0
10
8
6
4
2
0
-2
-4
-6
-8
-10
560
510
460
410
360
310
260

A. 2D presentation in CATT acoustics.


SHOEBOX
Direct Sound Pressure Level :
1000 Hz,
Used :
Lspk: S1, S2
- Speaker Data Not Authorized -
Interference :
Interference Sum
(Speaker Phase used)
Bandwidth :
0
Direct SPL [dB]
Max: 102
102
101
100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
82
Min: 77
(c) EASE 4.0 + / 11.02.2001 21:06:46 / ADA WAhnert

B. Narrow band presentation in EASE 4.0.

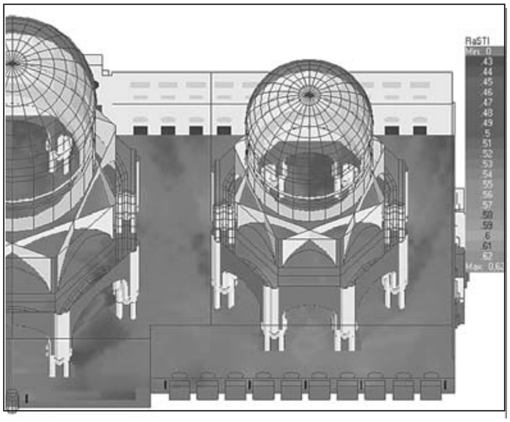

Ulysses Design Engine 2.30
125
250
500
1k
2k
4k
8k
125 Hz - 8000 Hz

C. Broad band presentation in ULYSSES 2.3.

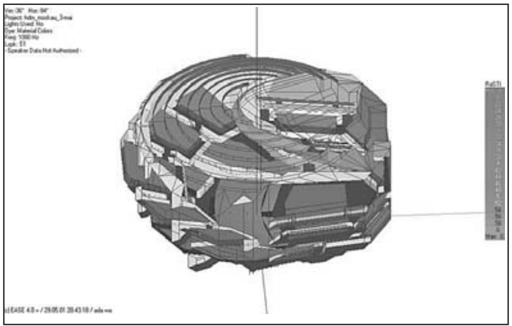
Figure 35-53. SPL mapping in simulation programs.

At first the unprocessed output of the DUT (raw data) is recorded and stored on the PC hard disk. Based on this original data set, the corresponding processing algorithm including band pass filters or time windows can be used multiple times with different parameters to look at the parts of interest.

A simple way to calculate the transfer function H(ω) from the recorded raw data is to divide the measured frequency response $Y(\omega)$ by the frequency response of the signal $X(\omega)$ (or by a reference response that was previously measured). The impulse response $h(t)$ can then be computed using the inverse Fourier Transform.


RaSTI
Min: 0
43
44
45
46
47
48
49
5
51
52
53
54
55
56
57
58
59
6
61
62
Max: 0,62

A. 3D presentation in a mosque.



B. 3D presentation in a concert hall.

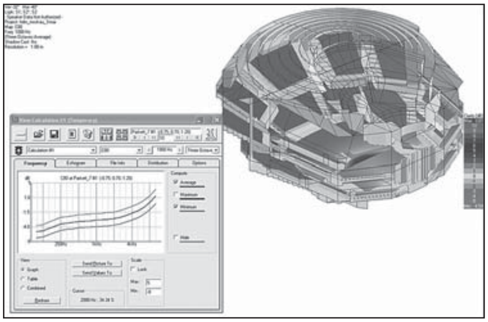
Figure 35-54. RaSTI presentations in simulation programs.

Until now it is common to utilize a static measuring procedure where the impulse response is derived in a separate step after every acoustic measurement. In contrast, a newly developed, dynamic method allows one to measure room acoustic impulse responses (RIR) in an efficient manner and to analyze this way the acoustic properties of an investigated acoustics space very user friendly—i.e., in real time,[57] see Fig. 35-57.

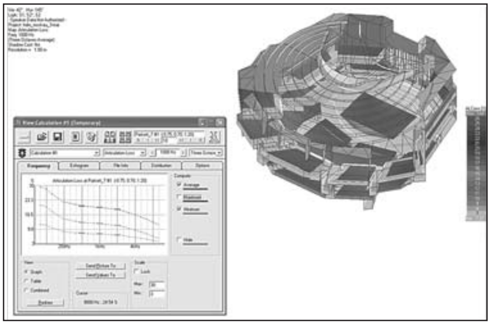
Determining the impulse response in real time means in this respect that gathering the acoustic source signals and calculating the impulse response data are a simultaneous and continuous process.

The dynamically derived RIR is because of a number of optimized postprocessing steps qualitatively absolutely equivalent to a statically derived RIR and may have typical lengths of 4–10 s.

The transformation between the frequency and time range is linear and of full length, analogous to the static procedure. Averaging can be likewise used to suppress the noise.



A. C80 presentation.



B. Articulation Loss ALCONS presentation.

Figure 35-55. AURA presentations in EASE 4.2.

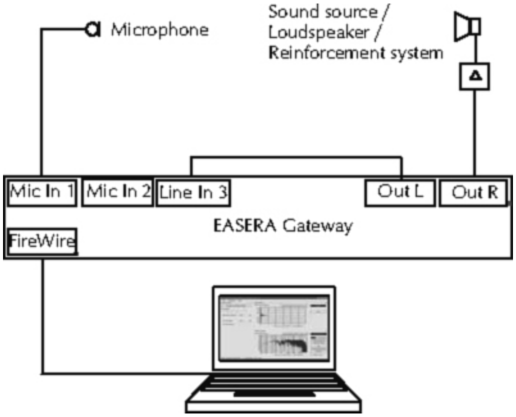

Microphone
Sound source /
Loudspeaker /
Reinforcement system
Mic In 1
Mic In 2
Line In 3
Out L
Out R
EASERA Gateway
FireWire

Figure 35-56. Block diagram of a modern software-based four-port measurement tool including the needed AD/DA convertor.

The real-time ability of the measuring system is based on very high refresh rates for the calculation of results and their display and analysis—approximately 10/s.

One can understand such a measuring system also as an "oscilloscope for room impulse responses." Possible changes of the acoustic behavior may be seen immediately and directly.

In Fig. 35-57, the excitation is done with noise, sweep, or MLS. In live situations this will be quite often annoying and cannot be done under all circumstances. So the next step is to use running music or speech signals as excitation signals and to derive impulse responses. Fig. 35-58 shows a block diagram for such a tool usable with natural signals like music or speech and in Fig. 35-59 the graphic user interface of EASERA SysTune[58] is shown, a tool that allows such a kind of measurements.

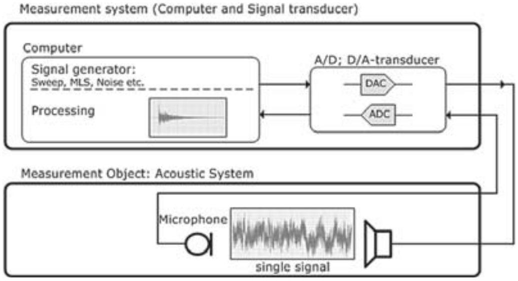

Measurement system (Computer and Signal transducer)
Computer
Signal generator:
Sweep, MLS, Noise etc.
Processing
A/D; D/A-transducer
DAC
ADC
Measurement Object: Acoustic System
Microphone
single signal

Figure 35-57. Static type of FFT-based measurements.

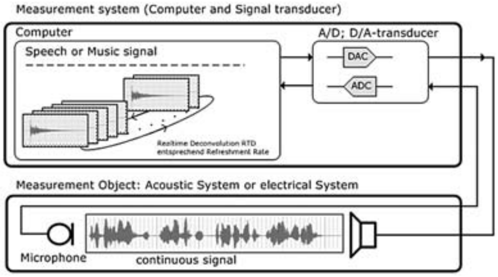

Measurement system (Computer and Signal transducer)
Computer
Speech or Music signal
A/D; D/A-transducer
DAC
ADC
Realtime Deconvolution RTD
entsprechend Refreshment Rate
Measurement Object: Acoustic System or electrical System
Microphone
continuous signal

Figure 35-58. Dynamic (continuous) type of FFT-based measurements.

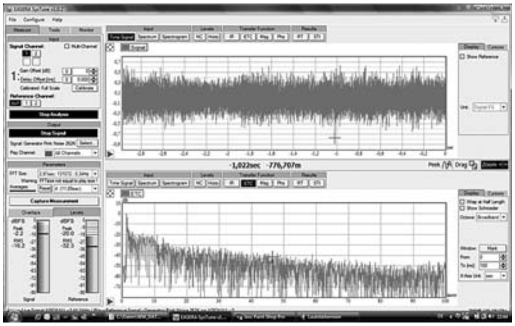


Figure 35-59. Graphic user interface of EASERA SysTune.

Once the IR has been computed in either a dynamic or static manner, electroacoustic and room-acoustic measures can be derived from it, such as the RT_{60}, D/R ratio, C50, or STI. These values can then be compared with the modeling results.

When performing such a comparison, it is always necessary to estimate the errors on each side, measurement and simulation, quantitatively in order to determine the significance of the deviations. The agreement between the results will depend on the degree to which measurement and model can provide reliable results.

References

1. *EASE 4.2 for Windows, Users Manual & Tutorial*, Renkus-Heinz Inc., March 2007.
2. Reichardt, W.; Kussev, A., “Ein- und Ausschwingvorgang von Musikinstrumenten und integrierte Hüllkurven ganzer Musikinstrumentengruppen eines Orchesters” (Initial and decay transients of complete instrumental groups and envelopes of an orchestra). Zeitschrift für elektrische Informations—und Energietechnik 3 (1973) 2, pp. 73–88.
3. *Taschenbuch der Akustik* (Handbook of acoustics) (Editor, Fasold, W., Kraak, W., Schirmer, W.), Verlag Technik, Berlin, 1984.
4. Tennhardt, H.-P., “Modellmessverfahren für Balanceuntersuchungen bei Musikdarbietungen am Beispiel der Projektierung des Großen Saales im Neuen Gewandhaus Leipzig” (Model measuring procedures for balance investigations of music performances demonstrated on the example of the great hall of the Neues Gewandhaus Leipzig). *Acustica* 56 (1984) Nr. 2 (Oktober 1984), pp. 126–135.
5. Knudsen, V. O., *Architectural Acoustics,* John Wiley and Sons, Inc., New York, 1932.
6. Meyer, J., *Akustik und musikalische Aufführungspraxi*s (Acoustics and music presentation practice), Verlag Erwin Bochinsky, Frankfurt am Main, 1999.
7. AESSC04-03, AES-2-R (specification of loudspeaker components used in professional audio and sound reinforcement) and AES04-01, AES-X83 (loudspeaker polar radiation measurements suitable for room acoustics).
8. Standard DIN 45570.
9. IEC Publication 268-5 Ed. 1972 “Sound system equipment part 5 loudspeaker, simulated program signal.”
10. Ahnert, W., Steffen, F., *Sound Reinforcement Engineering, Fundamentals and Practice*, London, E&FN Spon 2000, USA and Canada, New York, Routledge, 2000.
11. OIRT Empfehlung 55/1 Technische Parameter von Studio-Abhöreinrichtungen (Matinkylä 1985).
12. Stenzel, H., “Über die Richtwirkung von Schallstrahlern, Elektrische Nachrichten-Technik, Band 4” (1927), Seite 239.
13. Stenzel, H., *Leitfaden zur Berechnung von Schallvorgängen*, Verlag von Julius Springer, Berlin 1939.
14. Olson, H. F., *Elements of Acoustical Engineering*, D. van Nostrand Company, 2nd edition, New York, 1947.
15. Ureda, M. S., “Wave Field Synthesis with Horn Arrays,” 100th AES Convention Copenhagen/Denmark, May 1996, preprint No. 4144.
16. Benecke, H, Sawade, S., *Strahlergruppen in der Beschallungstechnik*, Sonderdruck Funkpraxis 1951.
17. Heil, C., “Sound Fields Radiated by Multiple Sound Sources Arrays,” 92nd AES Convention, Vienna, 1992 March 24–27, preprint No. 3269.
18. Duran-Audio BV, “Modelling the directivity of DSP controlled loudspeaker arrays,” White paper, Zaltbommel/Holland, June 2000, under www.duran-audio.nl.
19. Heinz, R., “DSP-Driven Vertical Arrays,” White paper, 2006, www.rh.com.
20. http://www.ateis-international.com/index.php.
21. Gunness, D., “Touring Line Array Technical Issues,” White paper, EAW, August 2000, under www.eaw.com.
22. Van Beuningen, G. W. J; Start, E. W., “Digital Directivity Synthesis of DSP Controlled Loudspeaker Arrays—A New Concept,” DAGA 2001, Hamburg-Harburg.
23. ISO 354: 2003, *Acoustics*—“Measurement of Sound Absorption in a Reverberation Room.”
24. Mommertz, E., “Determination of Scattering Coefficients from the Reflection Directivity of Architectural Surfaces,” *Applied Acoustics* 60 (2000) 201–203; see also, Vorländer, M., Mommertz, E., “Definition and Measurement of Random-Incidence Scattering Coefficients,” *Applied Acoustics* 60, 187–199, 2000.
25. ISO 17497-1: 2004 *Acoustics*—“Sound-Scattering Properties of Surfaces—Part 1: Measurement of the Random-Incidence Scattering Coefficient in a Reverberation Room.”
26. Bansal, M., “Wave Theory Based Extensions to Standard Room-Acoustic Particle Models,” PhD Thesis, Institute of Technical Acoustics, Technical University Berlin, 2008, to be published.

27. KEMAR-Head, http://www.gras.dk /00012/00330/.
28. EN 60268-4, "Sound system equipment. Microphones."
29. Kinsler, L., Frey, A., Coppens, A., Sanders, J., *Fundamentals of Acoustics*, 4th ed., Wiley, New York, 2000.
30. Ahnert, W., Feistel, S., "Cluster Design with EASE for Windows," presented at the 106th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.* (Abstracts), Vol. 47, p. 527 (1999 June), convention paper 4926.
31. Baird, J., Meyer, P., "The Analysis, Interaction, and Measurement of Loudspeaker Far-Field Polar Patterns," presented at the 106th Convention of the Audio Engineering Society, *J. Audio Eng. Soc.*, Vol. 47, June 1999.
32. Ahnert, W., Feistel, S., Baird, J., Meyer, P., "Accurate Electroacoustic Prediction Utilizing the Complex Frequency Response of Far-Field Polar Measurements," presented at the 108th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (*Abstracts*),* Vol. 48, p. 357 (2000 April), convention paper 5129.
33. Dynamic Link Library, http://msdn2.microsoft.com/en-us/library/ms682589.aspx.
34. Feistels., Ahnert, W., Bock, S., "New Data Format to Describe Complex Sound Sources," presented at the 119th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (*Abstracts*),* Vol. 53, pp. 1239, 1240 (2005 December), convention paper 6631, "GLL format specification,", http://www.ada-acousticdesign.de, http://www.sda.de.
35. Ahnert, W., Bourillet, C., Feistel, S., "Phase Presentation in the Acoustic Design Program EASE," presented at the 110th AES Convention, May 2001, special print under http://www.ada-acousticdesign.de.
36. CATT-Acoustic software, www.catt.se.
37. EASE software, http://www.ada-acousticdesign.de.
38. Feistel, S., Ahnert, W., "The Significance of Phase Data for the Acoustic Prediction of Combinations of Sound Sources," presented at the 119th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (*Abstracts*),* Vol. 53, p. 1240 (2005 December), convention paper 6632.
39. Feistel, S., Ahnert, W., "Modeling of Loudspeaker Systems Using High-Resolution Data," *J. Audio Eng. Soc.*, Vol. 55, pp. 571–597 (2007 July/August).
40. Anselm Goertz, www.anselmgoertz.de; Pat Brown, www.etcinc.us; Ron Sauro, www.nwaalabs.com.
41. Feistel, S., Ahnert, W., Hughes, C., Olson, B., "Simulating the Directivity Behavior of Loudspeakers with Crossover Filters," presented at the 123rd Convention of the Audio Engineering Society, New York City, NY, 2007 October 5–8, convention paper 7254.
42. Moldrzyk, C., Feistel, S, Ahnert, W., "Verfahren zur binauralen Wiedergabe akustischer Signale," Patent No. DE 10 2006 018 490.4.
43. AFMG Measurement Workshop, 123rd AES, New York, 2007.
44. Dalenbäck, B.-I., "Verification of Prediction Based on Randomized Tail-Corrected Cone-Tracing and Array Modeling," 137th ASA/2nd EAA Berlin (1999 March).
45. Van Maercke, (D. Martin,) J., "The Prediction of Echograms and Impulse Responses within the Epidaure Software," *Applied Acoustics,* Vol. 38, no. 2–4, p. 93 (1993).
46. Naylor, G. M., "Odeon—Another Hybrid Room Acoustical Model," *Applied Acoustics,* Vol. 38, no. 2–4, p. 131 (1993).
47. Farina, A., "Ramsete—A New Pyramid Tracer for Medium and Large Scale," *Proceedings of EURO-NOISE 95 Conference*, Lyon, 21–23 March 1995.
48. Vorländer, M., "Simulation of the Transient and Steady State Sound Propagation in Rooms Using a New Combined Sound Particle–Image Source Algorithm," *J. Acoust. Soc. Am.*, 86, 172.
49. Schmitz, O.; Feistel, S.; Ahnert, W.; Vorländer, M., "Merging Software for Sound Reinforcement Systems and for Room Acoustics," presented at the 110th AES Convention (2001) May 12–15, Amsterdam, preprint No. 5352.
50. Rubin, S. M., Whitted, T., *Computer Graphics,* Vol. 14, no. 3, 1980.
51. Fujimoto, A, Tanaka, T., Iwata, K., *IEEE Computer Graphics and Applications,* Vol. 6, no. 4, 1986.
52. Havran, V., Prikryl, J., Purgathover, W., Tech. Rep. TR-186-2-00-14, Institute of Computer Graphics, Vienna University of Technology, 2000.
53. Amanatides, J., Woo, A., Eurographics '87 Conference Proceedings (1987).
54. Feistel, S., Ahnert, W.; Miron, A., Schmitz, O., "Improved Methods for Calculating Impulse Responses with EASE 4.2 AURA," 19th International Congress on Acoustics, September 2007, Madrid.
55. ODEON software, version 9.1, www.odeon.dk.

56. Ureda, M. S., “Line Arrays, Theory and Applications,” 110th AES Convention, March 12—15 Amsterdam 2001, preprint No. 5304.
57. EASERA software, http://www.Easera.com.
58. EASERA SysTune software, http://www.EaseraSysTune.com.

Chapter 36

Designing for Speech Intelligibility

by Peter Mapp

36.1 Introduction 1387
36.2 Parameters Affecting Speech Intelligibility 1387
36.3 The Nature of Speech 1387
36.4 Factors Affecting Sound System Intelligibility 1388
  36.4.1 Primary Factors Include 1388
  36.4.2 Secondary Factors Include 1388
36.5 System Frequency Response and Bandwidth 1388
36.6 Loudness and Signal-to-Noise Ratio 1390
36.7 Reverberation Time and Direct-to- Reverberant Ratios 1393
  36.7.1 Intelligibility Prediction—Statistical Methods 1395
  36.7.2 Intelligibility and Reverberation Time 1397
36.8 Some Further Effects of Echoes and Late Reflections 1397
36.9 Uniformity of Coverage 1399
36.10 Computer Modeling and Intelligibility Prediction 1400
36.11 Equalization 1400
36.12 Talker Articulation and Rate of Delivery 1401
36.13 Summary of Intelligibility Optimization Techniques 1402
36.14 Intelligibility Criteria and Measurement 1403
  36.14.1 Subject-Based Measures and Techniques 1403
  36.14.2 Objective Measures and Techniques 1404
    36.14.2.1 Articulation Index 1404
    36.14.2.2 Articulation Loss of Consonants 1404
    36.14.2.3 Direct-to-Reverberant and Early-to-Late Ratios 1404
    36.14.2.4 Speech Transmission Index STI, RASTI, and STIPA 1405
    36.14.2.5 SII Speech Intelligibility Index 1409
  36.14.3 The Future for Speech Intelligibility Measurements 1409
Acknowledgments 1410
Bibliography 1410

This page intentionally left blank

36.1 Introduction

The fundamental purpose of a paging, announcement, voice alarm, or speech reinforcement system is to deliver intelligible speech to the listener. A surprising number of systems, however, fail to achieve this basic goal. There can be many reasons for this, ranging from inadequate signal-to-noise ratio to poor room acoustics or inappropriate choice or location of the loudspeaker. It is the job of the sound system designer to be aware of these factors and take them into account when designing a sound system and selecting devices to provide the degree of intelligibility required. In order to do this, however, an understanding of the basic factors that affect speech intelligibility and the way we hear speech is required. This chapter therefore begins by taking a look at the nature of the speech signal and how we hear it before discussing design strategies and ways of optimizing system design and performance. Current methods of assessing and measuring intelligibility are then also discussed together with comments on their practical limitations.

36.2 Parameters Affecting Speech Intelligibility

Although sound quality and speech intelligibility are inextricably linked, they are not the same thing. For example it is quite possible to have a poor sounding system that is highly intelligible (e.g., the frequency response limited and resonant re-entrant horn) or alternatively a high-quality system that is virtually unintelligible (e.g., a hi-fi loudspeaker in an aircraft hangar). Similarly a common mistake, often made when discussing intelligibility, is to confuse audibility with clarity. Just because a sound is audible does not mean to say that it is intelligible. Audibility relates to the ability of a listener to physically be able to hear a sound, whereas clarity describes the ability to detect the structure of the sound. In the case of speech, this means hearing the consonants and vowels correctly in order to identify the words and sentence structure and so give the speech sounds intelligible meaning.

36.3 The Nature of Speech

A speech signal involves the dimensions of sound pressure, time, and frequency. Fig. 36-1 shows some typical speech waveforms representing the numbers "one" "two," and "three." The waveforms are highly complex, with amplitudes and frequency contents that change almost millisecond by millisecond. Consonant sounds typically have durations of around 65 ms and vowels 100 ms. The duration of syllables is typically 300–400 ms whereas complete words are about 600–900 ms in length dependent on their complexity and rate of speech. When speech is transmitted into a reverberant space, local reflections and the general reverberation distort the speech waveform by smearing it in time. The reverberant tail of one syllable or word can overhang the start of the next and so mask it, thereby reducing the potential clarity and intelligibility, Fig. 36-2. Equally if the background noise level is high or more accurately if the speech signal-to-noise ratio is too low, then again parts of words or syllables become lost and intelligibility deteriorates, Fig. 36-3. There are many other factors that can affect the potential intelligibility and perceived clarity of a speech signal, the most important are summarized below.

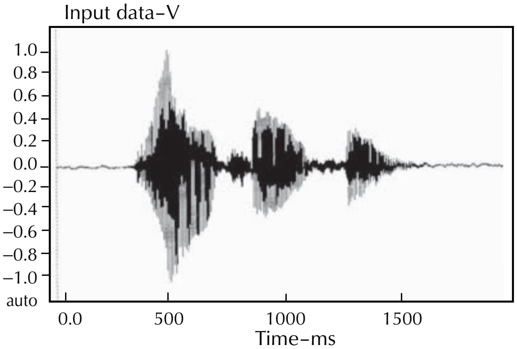


Figure 36-1. Anechoic speech waveforms for the numbers "one," "two," and "three."

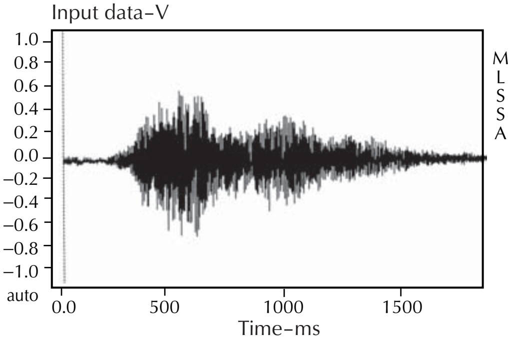


Figure 36-2. Speech waveforms (as Fig. 36-1) but with reverberation (RT_{60} = 2.4 s). The way one word runs into the next can clearly be seen, but with concentration the individual words can still be understood.

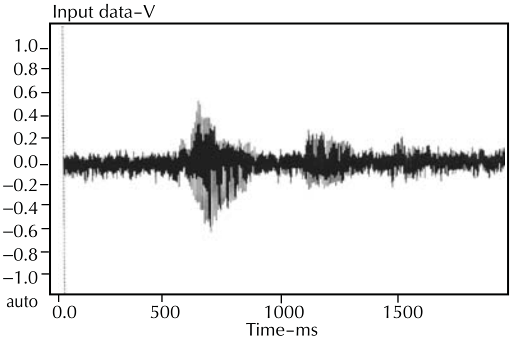


Figure 36-3. Speech waveforms (as Fig. 36-1) in background noise. The noise masks much of the waveform detail but the speech remains intelligible.

36.4 Factors Affecting Sound System Intelligibility

36.4.1 Primary Factors Include

- Sound system bandwidth and frequency response.
- Loudness and signal-to-noise ratio.
- Room reverberation time (RT_{60}).
- Volume and size and shape of the space.
- Distance from the listener to a loudspeaker.
- Directivity of the loudspeaker.
- The number of loudspeakers operating within the space.
- The direct to reverberant ratio** (this is directly dependent upon the previous 5 factors).
  - * Talker annunciation/rate of delivery.
  - * Listener acuity.

** Strictly speaking a more complex characteristic than the simple D/R ratio should be used. Better correlation with perceived intelligibility is obtained by using the ratio of the direct sound and early reflected energy to late reflected sound energy and reverberation. This may be termed C50 or C35 depending upon the split time used to delineate between the useful and deleterious sound arrivals.

36.4.2 Secondary Factors Include

- System distortion (e.g., harmonic or intermodulation).
- System equalization.
- Uniformity of coverage.
- Presence of very early reflections (<1–2 ms).
- Sound focusing or presence of late or isolated higher-level reflections (>70 ms).
- Direction of sound arriving at the listener.
- Direction of any interfering noise.
  - * Gender of talker.
  - * Vocabulary and context of speech information.
  - * Talker microphone technique.

The bulleted parameters marked with a bullet (•) are building or system related, while those marked with an asterik (*) relate to human factors outside the direct control of the system itself.

How each of the above factors affects the potential intelligibility of a sound system is discussed below together with ways that a system designer can minimize the deleterious effects and optimize the desirable characteristics.

36.5 System Frequency Response and Bandwidth

Speech covers the frequency range from approximately 100 Hz–8 kHz, although there are also higher harmonics affecting the overall sound quality and timbre extending up to 12 kHz and above. Fig. 36-4 shows an averaged speech spectrum with the relative frequency contributions in octave bands. Maximum speech energy occurs over the approximate range 200–600 Hz—i.e., in the 250 Hz and 500 Hz octave bands, and falls off rapidly at about 6 dB per octave at higher frequencies as can be seen in Fig. 36-4.

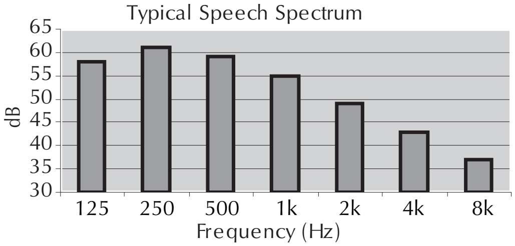


Figure 36-4. Average speech spectrum (octave band resolution).

The lower frequencies correspond to the vowel sounds whereas the weaker upper frequencies correspond to the consonants. The contributions to speech intelligibility, however, do not follow the same pattern—indeed quite the reverse. Fig. 36-5 shows the relative octave band percentage contributions to intelligibility. Here we can clearly see that most intelligibility is concentrated in the 2 kHz and 4 kHz bands, these contributing approximately 30% and 25%, respectively, while the 1 kHz octave contributes a further 20%. These three bands therefore provide over 75% of the available spectral intelligibility content.

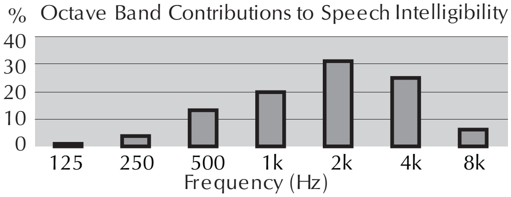


Figure 36-5. Octave band percentage contributions to speech intelligibility.

Whereas the range 300–3000 Hz has been shown to be adequate for telephone intelligibility, a wider range is generally required for sound system use—particularly under more difficult acoustic conditions. This effect is shown in Fig. 36-6. This contrasts the results of telephone (monophonic listening) with some recent research carried out by the author in a reverberant space (RT_{60} = 1.5 s). The upper curve after Fletcher (1929) shows that the contribution to intelligibility hardly increases beyond 4 kHz, while the lower curve, made on a system in a real space (binaurally) shows improvements occurring up to 10 kHz. The need for an extended bandwidth can therefore immediately be seen. Limited bandwidth should not be a problem with modern sound system equipment and loudspeakers. However, there are some notable exceptions. These include:

1. Inexpensive poor-quality microphones.
2. Some re-entrant horn loudspeakers (or CD horn drivers used without equalization).
3. Some inexpensive digital message stores.
4. Miniature, special purpose loudspeakers.

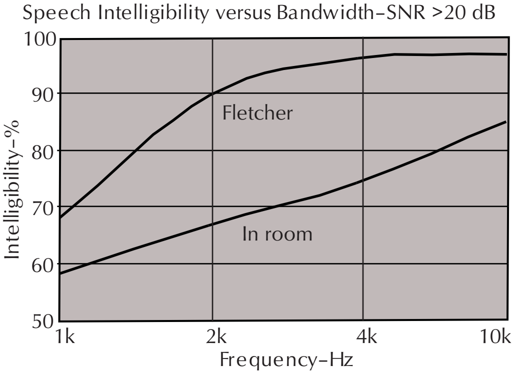


Figure 36-6. Effect of frequency bandwidth on speech intelligibility. Upper curve—monophonic listening (after Fletcher). Lower curve—binaural listening (after Mapp).

Many potentially adequate sound systems are often let down by employing a cheap or restricted bandwidth microphone at the front end of the system. In the author's experience, even on a basic paging system employing restricted bandwidth loudspeakers—e.g., re-entrant horns—the difference between a microphone with a reasonably wide and well-controlled frequency response can always be readily identified over one with a restricted response, even if it exceeds the response of the loudspeakers themselves. Rubbish in equals rubbish out is certainly the case here. However, when operating under high-background noise conditions, a compromise may need to be reached between optimal frequency response and optimal noise rejection, as the two parameters are often divergent.

Apart from component equalization (or the lack of it) by far the most common problems associated with system frequency response stem from either loudspeaker/boundary-room effects or interactions between closely spaced (multiple) loudspeakers. Fig. 36-7 shows the effect of positioning a high-quality monitor loudspeaker with an impeccably flat response close to a boundary wall. As can be seen the response is now far from flat!

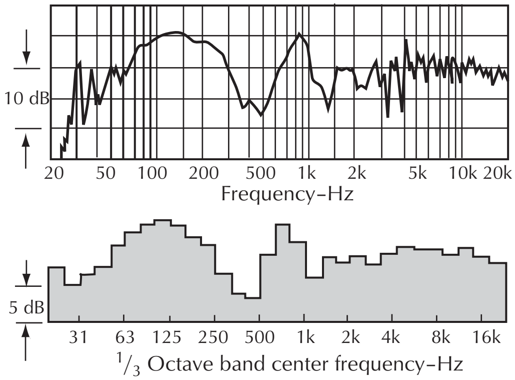
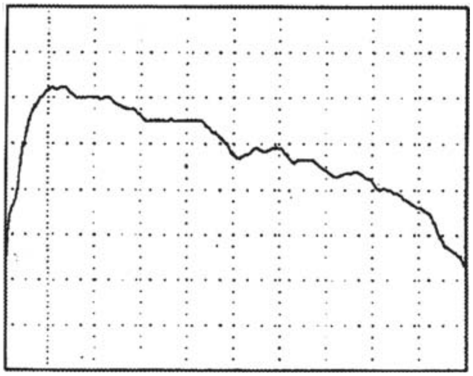


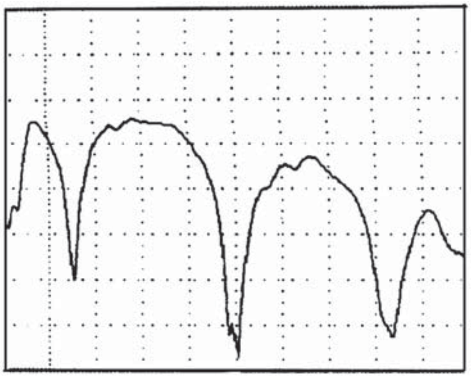
Figure 36-7. Effect of local boundary interaction on loudspeaker frequency response.

Equalization alone cannot correct for this problem. Reduction of the peaks is possible but the notches in the response cannot be equalized out as they are caused by complex phase interactions that cannot be corrected by means of frequency filtering. Interaction between loudspeakers is a common problem in cluster design, where the radiated wave fronts can suffer from missynchronization due to different acoustic path lengths occurring—e.g., due to differences between acoustic centers. Fig. 36-8 shows a typical interaction problem (after Davis and Davis).



Vertical: 6 dB/division
Horizontal: 50.33 Hz – 10,001.20 Hz
Resolution: 5.3674E + 02 Hz
Both horns: Near throw delayed 300 µs

A. Loudspeakers in synchronization.

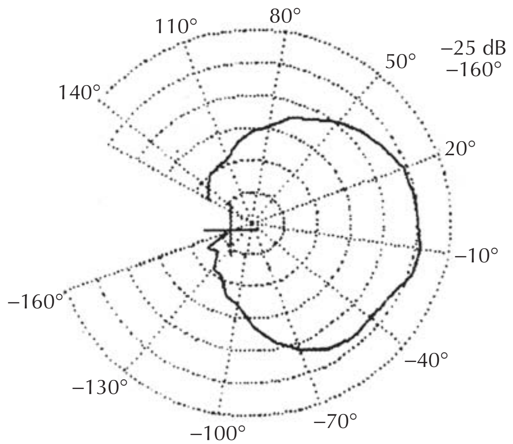


Vertical: 6 dB/division
Horizontal: 50.33 Hz – 10,001.20 Hz
Resolution: 5.3674E + 01 Hz
Both horns: No delay

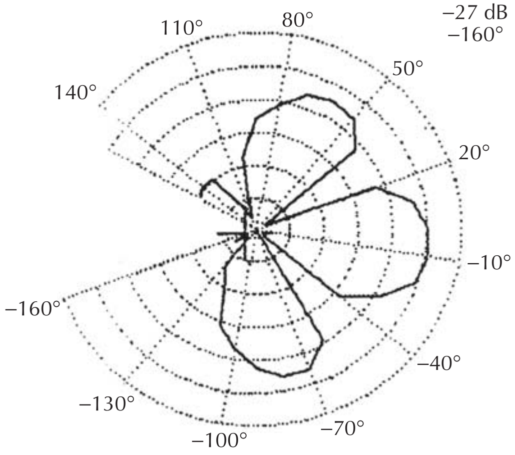
B. Loudspeakers out of synchronization by 300 µs.

Figure 36-8. Frequency response of two loudspeakers. Upper curve shows effect when the sound arrivals are synchronized, lower curve shows effect of 300 µs missynchronization.

Here the frequency response of two horn loudspeakers is shown. In the upper curve, the sound arrivals are synchronized and hence add constructively. However, in the lower curve, the horns are missynchronized by just 300 µs. A series of sharp comb filters occur. Not only is useful speech information lost in the extensive series of nulls but also the polar radiation pattern is often undesirably affected as shown in Fig. 36-9. The resultant lobes may not only result in certain frequencies not being transmitted to the listeners, but lobes may also be created that can cause undesirable reflections to occur. These may cause either additional unwanted excitation of the reverberant field or cause the generation of late reflections (echoes) that may damage intelligibility.



A. Two loudspeakers in synchronization.



B. Two loudspeakers out of synchronization.

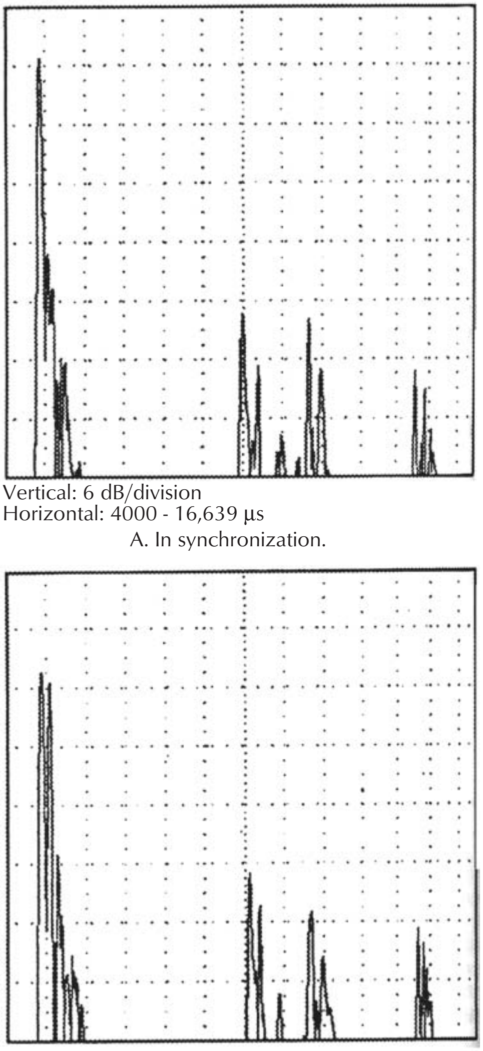
Figure 36-9. Polar response of two loudspeakers in and out of synchronization.

Fig. 36-10 shows the corresponding ETC reflection sequence for the horns in a reverberant space in and out of synchronization. Note the increased excitation of the reverberant field when out of synchronization.

The lobes caused my missynchronization, apart from potentially reducing intelligibility, may also reduce system feedback margin, either by directly radiating sound back to a live microphone or by causing a strong early reflection to occur back into the microphone.

36.6 Loudness and Signal-to-Noise Ratio

The sound level produced by a sound system must be adequate for the intended listeners to be able to hear it comfortably. If the level is too low, many people, partic-



A. In synchronization.

B. Out of synchronization.

Figure 36-10. ETC curves of the two loudspeakers shown in Fig. 36-9.

ularly the elderly or those suffering even a mild hearing loss, may miss certain words or strain to hear, even under quiet conditions. Although normal face to face conversation may take place at around 60 dBA, regularly listeners demand higher sound pressure levels from sound systems, with 70–75 dBA being typical for conference systems even when under quiet listening conditions.

In noisy situations, it is essential that a good SNR is achieved. Various rules of thumb have been developed over the years. As a general minimum, 6 dBA is required and at least 10 dBA should be aimed for. Above 15 dBA there is still some improvement to be had, but the law of diminishing returns sets in for most practical systems.

There is some disagreement among the generally accepted reference data. Fig. 36-11, for example, shows the general relationship between SNR and intelligibility. As the curve shows, this is an essentially linear relationship. In practice, the improvement curve flattens out at high signal-to-noise ratios—though this is highly dependent on the test conditions. This fact is shown in Fig. 36-12, which compares the results of a number of studies, using different test conditions and signals.

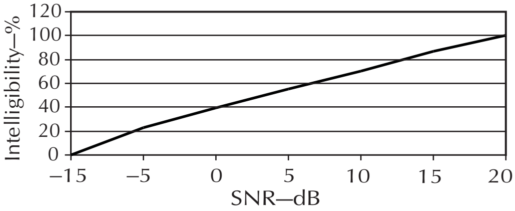


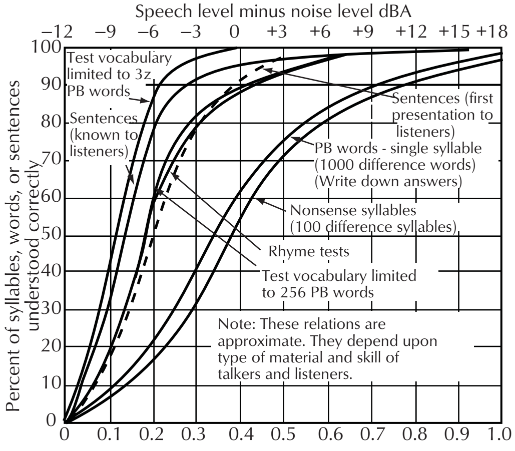
Figure 36-11. Effect of SNR on speech intelligibility.

The curve, for example, shows that for more difficult listening tasks, the greater the SNR has to be in order to achieve good intelligibility. Fig. 36-13 shows the effect of SNR on the %Alcons intelligibility scale. Here, the improvement can be seen to clearly flatten out above 25 dB SNR. Under high noise conditions, such a SNR could demand excessively high SPLs and caution must be exercised.

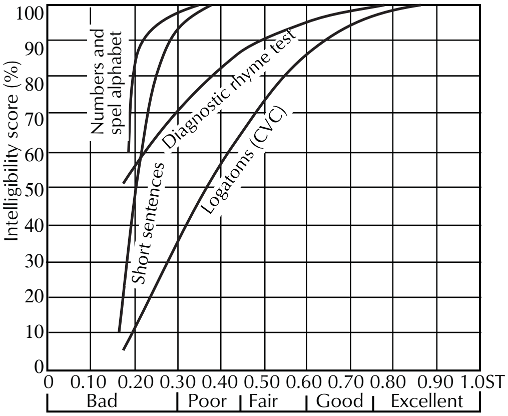
Where noise is a particular problem, a full spectral analysis should be carried out. Ideally this should be in terms of ⅓ octaves but for many applications ¹⁄₁ octave band analysis will be adequate and certainly more informative than a single dBA value. Fig. 36-14 shows such an analysis.

In the upper curve, which depicts a positive SNR, it can be seen that the speech signal is greater than the noise over each of the octave band frequencies. However, in the lower curve, it can be seen that at high frequencies the noise exceeds the desired speech signal. The overall effect on potential intelligibility can be calculated by looking at the individual octave band SNRs and then weighting and summing them in accordance to their relative contributions as shown earlier in Fig. 36-5.

This is the basis of the Articulation Index (AI), which is a good measure for determining the effects of noise on speech—either over single channel transmission lines such as telephone or radio communications or over PA



A. Articulation Index (AI)



B. Speech Transmission Index (STI)

Figure 36-12. Comparison of speech intelligibility test formats as a function of Articulation Index (AI) and Speech Transmission Index (STI).

systems in low but noisy spaces. The AI method is not able to take account of room reverberation or reflections.

In many situations, the background noise may not be steady but vary over time. This is particularly the case in many industrial complexes or transportation concourses. Spectator sports also can exhibit highly variable crowd noise levels dependent on the action at any given time. Fig. 36-15 shows a typical noise profile for an underground train station. Peaks of 90 dBA plus were recorded as the trains moved in and out of the platforms. A PA system would therefore need to generate at least 96–100 dBA in order to achieve an appropriate SNR at these times.

Noise sensing and automatic level control are essential under such conditions, otherwise during the relatively quiet periods when ambient levels drop down to around only 66 dBA, significant startle may be caused by such high-level announcements. (A better solution is to store announcements and wait for the regularly occurring quieter periods rather than trying to compete with the background noise all the time.)

Spectator sports can also create wildly fluctuating noise levels. Again if possible, announcements should be made during the quieter periods, the levels of which can be best determined by a statistical analysis of the crowd behavior at the particular venue in question. Fig. 36-16 shows part of the time history for a soccer match. Note that peak values in excess of 110 dBA can occur.

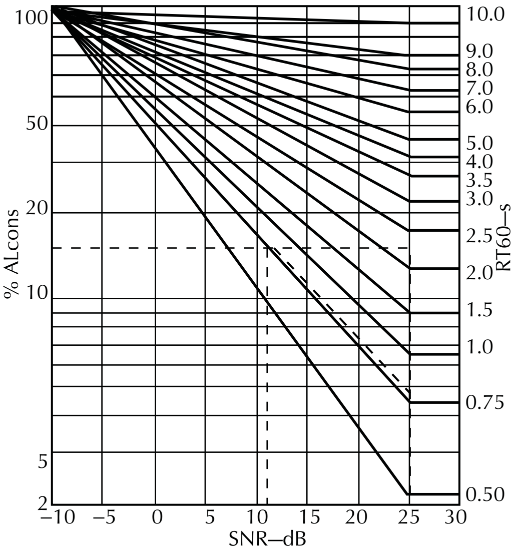
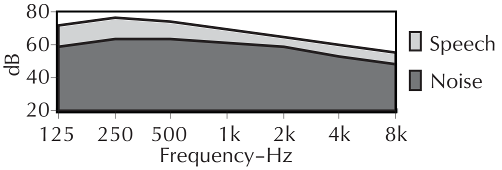
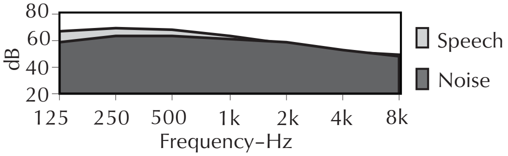


Figure 36-13. Effect of signal to noise ratio on %*Alcons*.

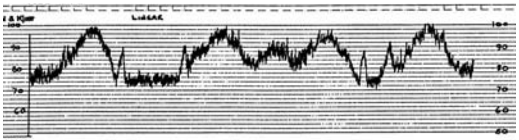


A. Good SNR. Good intelligibility—noise does not mask speech.

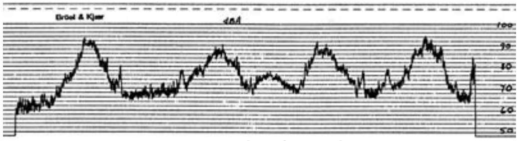


B. Poor SNR. Poor intelligibility—note complete masking of high frequency speech sounds by the background noise.

Figure 36-14. Spectral analysis of speech-to-noise ratio (one octave band resolution).



A. Broadband linear recording.



B. "A" weighted time history.

Figure 36-15. Noise-time history profile of underground trains entering and leaving station.

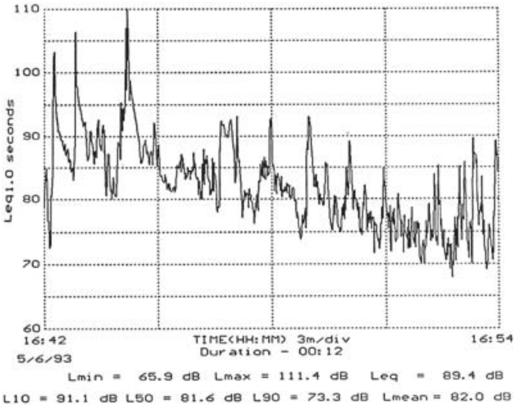


Figure 36-16. Noise-time history analysis of soccer game crowd noise—note short-term variability and peak of 111 dB as compared to average of 82 dB.

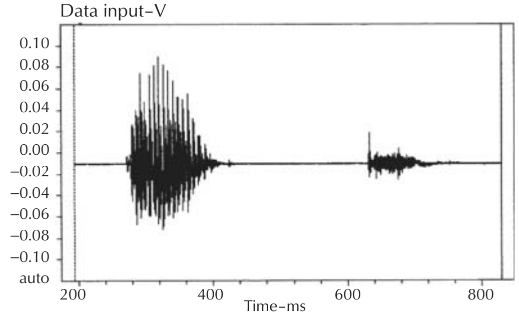
It must not be forgotten that any noise occurring at the microphone itself will reduce the perceived SNR—indeed this is directly additive to the SNR at the listener's position. At least 20 dBA should be aimed for and preferably >25 dBA. A number of techniques can be employed to achieve this, including:

- Close talking/noise canceling microphones.
- Use of highly directional microphones (e.g., gun microphones or adaptive arrays).
- Providing a noise hood or preferably by locating the microphone in a suitable quiet room or enclosure.
- Digital, noise canceling and processing can also be used in extreme conditions to improve the SNR.

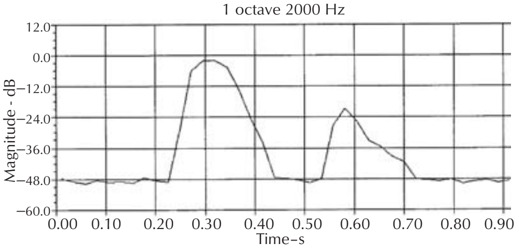
36.7 Reverberation Time and Direct-to-Reverberant Ratios

Just as noise can mask speech signals so too can excessive reverberation. However, unlike the simpler case of SNR, the way in which the direct-to-reverberant (D/R) ratio affects speech intelligibility is not constant but depends on the room reverberation time, the level of the reverberant sound field and on the nature of the speech itself.

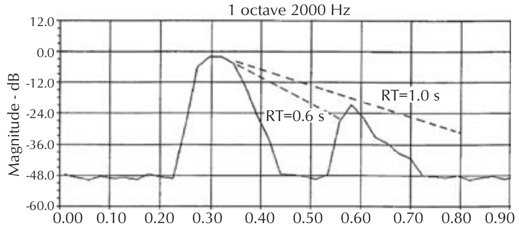
The effect is illustrated in Fig. 36-17A–C. The upper trace is the speech waveform of the word *back.* The word starts suddenly with the relatively loud "ba" sound. This is followed some 300 ms later by the consonant "ck" sound. Typically the "ck" sound will be 20 dB – 25 dB lower in amplitude than the "ba" sound.



A. Basic waveform.



B. Temporal amplitude response–dB.



C. Temporal amplitude response with effects if reverberation time is superimposed.

Figure 36-17. Waveform of the word *back*

With short reverberation times—e.g., 0.6 s—the "ba" sound has time to die away before the start of the "ck" sound. Assuming a 300 ms gap, the "ba" will have decayed by around 30 dB and will not mask the later "ck." However, if the reverberation time increases to 1 s and if the reverberant level in the room is sufficiently high (i.e., a low *Q* device is used), then the “ba” sound will have only decayed by approximately 18 dB and will completely mask the “ck” sound by 8 dB to 13 dB. It will therefore not be possible to understand the word *back* or distinguish it from similar words such as *bat, bad, bath,* or *bass* since the all important consonant region will be lost. However, when used in the context of a sentence or phrase, the word may well be worked out by the listener from the context. Further increasing the reverberation time (or reverberant level) will further increase the degree of masking.

Not all reverberation, however, should necessarily be considered to be a bad thing, a degree of reverberation is essential to aid speech transmission and to aid the talker by returning some of the sound energy back to him or her. This enables subconscious self-monitoring of their speech signal to occur and so feed back information about the room and projected level. The room reverberation and early reflections will not only increase the perceived loudness of the speech, thereby acting to reduce the vocal effort and potential fatigue for the talker, but also provide a more subjectively acceptable atmosphere for the listeners. (No one would want to live in an anechoic chamber.) However, as we have seen the balance between too much or not enough reverberation is a relatively fine one.

The sound field in a large space can be highly complex. Statistically, it can be divided into two basic components, the direct field and the reverberant field. However, from the point of view of subjective impression and speech intelligibility the sound field needs to be further subdivided to produce four distinct components. These are:

1. Direct Sound—that directly from source to listener.
2. Early Reflections—arriving at the listener approximately 35–50 ms.
3. Late Reflections—arriving at the listener approximately 50–100 ms later (though discrete reflections can also be later than this).
4. Reverberation—high density of reflections arriving after approximately 100 ms.

Fig. 36-18 summarizes the sound field components discussed above.

To the above list one could also add “Early Early” reflections—those occurring within 1–5 ms. (If specular in nature, these generally cause comb filtering and sound coloration to occur. Reflections of 1–2 ms are particularly troublesome as they can cause deep notches in the frequency response to occur around 2 kHz and thereby reduce intelligibility by attenuating the primary speech intelligibility frequency region.)

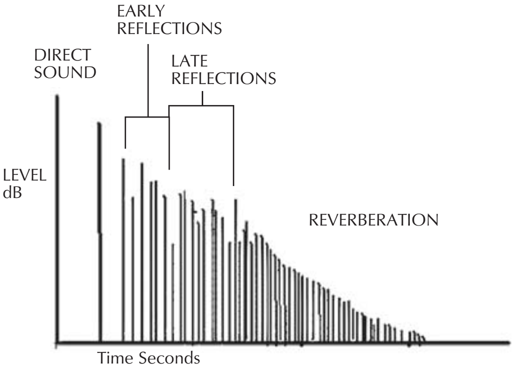


Figure 36-18. Sound field components.

Opinion as to how the direct sound and early reflections integrate is currently somewhat divided. Many believe that reflections occurring up to around 35–50 ms after the direct sound fully integrates with it, provided that they have a similar spectrum. This causes an increase in perceived loudness to occur, which under noisy conditions can increase the effective SNR and hence intelligibility. Under quieter listening conditions, however, the case is not quite so clear, with factors including spectral content and direction of reflection becoming increasing important. Equally some research suggests that the integration time may be frequency dependent but generally around 35 ms for speech signals. However, there is general agreement that later arriving reflections (>50 ms) act such as to degrade intelligibility with increasing effect as the arrival time delay increases.

Sound arriving after approximately 100 ms generally signals the start of the reverberant field though strong discrete reflections arriving after 60 ms or so will be heard as discrete echoes. It is the ratio of direct + early reflections to late reflections and reverberation that determines the potential intelligibility in a reverberant space (assuming that other effects such as background noise and frequency response considerations are neglected). As a rule, positive ratios are desirable but rarely achieved in reality, though there are exceptions.

This is demonstrated in Figs. 36-19 and 36-20. Fig. 36-19 shows the energy time curve (ETC) sound arrival analysis for a highly directional (high *Q*) loudspeaker in a large reverberant church (RT_{60} = 2.7 s at 2 kHz). The D/R ratio at the measuring position (approximately $^2/_3$ way back) is 8.7 dB resulting in a high degree of intelligibility. Other intelligibility measures taken from the same TEF data (see Section 36-13 on measuring intelligibility) are:

- *%Alcons* 4.2%.
- Equivalent rasti 0.68.
- C50 9.9 dB.

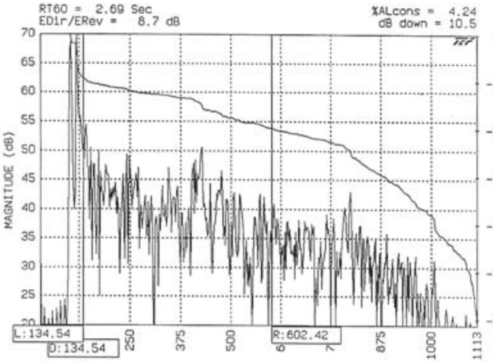


Figure 36-19. ETC of a high *Q* (highly directional) loudspeaker in reverberant church.

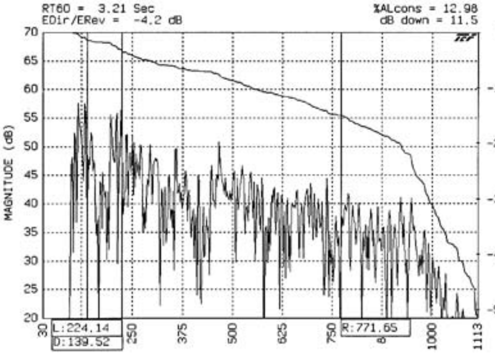


Figure 36-20. ETC of low *Q* (omnidirectional) loudspeaker in a reverberant church.

An opportunity to exchange the high *Q* device for an almost omnidirectional, low *Q* loudspeaker was taken and found to have a profound effect on the perceived intelligibility and the resulting ETC. This is shown in Fig. 36-20, which presents an obviously very different curve and pattern of sound arrivals. Clearly there is far more excitation of the reflected and reverberant sound fields. The D/R ratio is now –4 dB (a degradation of some 12 dB) and other computed data is:

- *%Alcons* is now only 13%.
- C50 has been reduced to –3.6 dB.
- Equivalent rasti to 0.48.

All the indicators and even a visual inspection of the graphs show there to be a significant reduction in the potential intelligibility.

While visual inspection of an ETC can be very enlightening, it can also at times be misleading. Take for example the curve shown in Fig. 36-21. At first glance this resembles the ETC for the low Q device shown above and might suggest low intelligibility since no clear direct sound component is visible. However, densely distributed ceiling loudspeaker systems in a controlled environment do not work in the same way as point source systems in large spaces. In the former, the object is to provide a dense, short path length sound arrival sequence, from multiple nearby sources. The early reflection density will be high and in well-controlled rooms, the later arriving reflections and reverberant field will be attenuated. This results in smooth coverage and high intelligibility. In the case shown in Fig. 36-21, the RT_{60} was 1.2 s and the resulting C50 was +2.6 dB and the Rasti was 0.68, results both indicating high intelligibility, which indeed was the case.

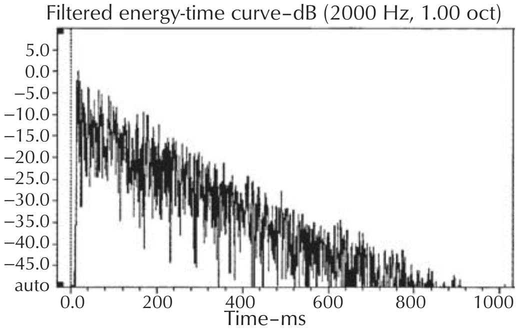


Figure 36-21. ETC of distributed ceiling loudspeaker system in an acoustically well-controlled room.

It should not be forgotten that whereas it may well be possible to produce high intelligibility in a localized area, even in a highly reverberant space, extending the coverage to a greater area will always result in reduced intelligibility at this point, as the number of required sources (additional loudspeakers) to accomplish the task increases. This is primarily due to the resulting increase in acoustic power fed into the reverberant field (i.e., increase in reverberant sound level) often referred to as the loudspeaker system *n* factor.

36.7.1 Intelligibility Prediction—Statistical Methods

While it is relatively trivial to accurately calculate the direct and reverberant sound field components by means of traditional statistical acoustics, it is not possible to accurately estimate, on a statistical basis, the early and late reflection fields. (To do this requires a computer model of the space and ray-tracing/reflection analysis program.)

Prior to such techniques being available, a number of statistically based intelligibility prediction methods based on calculation of the direct and reverberant fields were developed and are still useful in order to provide a quick ball park review of a design or idea. They have greater accuracy when applied to center cluster or point source systems as opposed to distributed loudspeaker systems (particularly high-density distributed systems).

The best known equation is that of Peutz as later modified by Klein and is the articulation loss of consonants equation (%*Alcons*). Peutz related intelligibility to a loss of information. For a loudspeaker-based system in a reverberant room, the following factors are involved:

- Loudspeaker directivity (Q).
- Quantity of loudspeakers operating in the space (n).
- Reverberation time (RT_{60}).
- Distance between listener and loudspeaker (D).
- Volume of the space (V).

$$\%Alcons = \frac{200^{*}D^{2}(T_{60}^{\;2})(n+1)}{QV} \qquad (36\text{-}1)$$

* use 656 for American units

The %Alcons scale is unusual in that the smaller the number, the better the intelligibility. From Eq. 36-1 it can be seen that the intelligibility in a reverberant space is in fact proportional to the *volume* of the space and the *directivity* (Q) of the loudspeaker, (i.e., increasing either of these parameters while maintaining the others constant will improve the intelligibility). From the equation it can also be seen that intelligibility is inversely proportional to the squares of *reverberation time* and *distance* between the listener and the loudspeaker.

The equation was subsequently modified to take account of talker articulation and the effect that an absorbing surface has on the area covered by the loudspeakers.

$$Alcons = \frac{200^{*}D^{2}(T_{60}^{\;2})(n+1)}{QVma} + K \qquad (36\text{-}2)$$

* use 656 for American units

where,

m is the critical distance modifier, taking into account higher than average absorption of the floor with an audience, for example,

m is $(1-a)/(1-ac)$ where a is the average absorption coefficient, ac is the absorption in the area covered by the loudspeaker,

k is the listener/talker correction constant typically 1–3, but for poor listeners/talkers can increase to 12.5%.

Peutz found that the limit for successful communication was around 15% Alcons. From 10 to 5% intelligibility is generally rated as good and below 5% the intelligibility can be regarded as excellent. A limiting condition

$$Alcons = 9T + k \qquad (36\text{-}3)$$

was also found to occur by Peutz.

Although not immediately obvious from the equations, they are effectively calculating the direct-to-reverberant ratio. By rearranging the equation, the effect of the direct-to-reverberant ratio on %A*lcons* can be plotted with respect to reverberation time. This is shown in Fig. 36-21. From the figure, the potential intelligibility can be directly read from the graph as a function of D/R and rverberation time. (By reference to Fig. 36-13 the effect of background noise SNR can also be incorporated.)

The Peutz equations assume that the octave band centered at 2 kHz is the most important in determining intelligibility and uses the values for the direct level, reverberation time, and Q to be measured in this band. There is also an assumption that there are no audible echoes and that the room or space supports a statistical sound field being free of other acoustic anomalies such as sound focusing.

In the mid-1980s Peutz redefined the %*Alcons* equations and presented them in terms of direct and reverberant levels and background noise level.

$$\%Alcons = 100(10^{-2(A+BC)-ABC}) + 0.015 \qquad (36\text{-}4)$$

where,

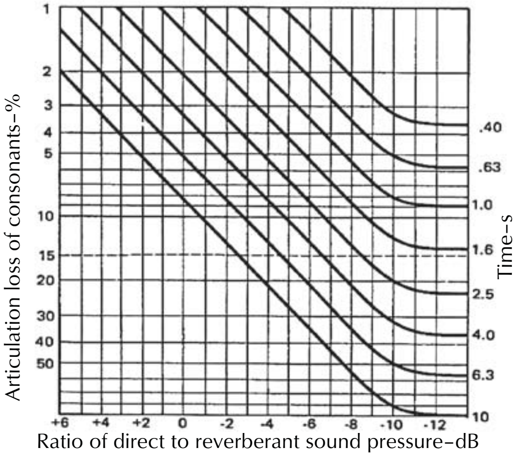
$$A = -0.32\log\left[\frac{L_R + L_N}{10L_D + L_R + L_N}\right]$$

for $A \geq 1$, let $A = 1$

$$B = -0.32\log\left[\frac{L_N}{10L_R + L_N}\right]$$

for $B \geq 1$, let $B = 1$

$$C = -0.50\log\left(\frac{RT_{60}}{12}\right)$$



Probable articulation loss of consonants versus reverberation time and direct-to-reverberant sound ratio.

Figure 36-22. Effect of direct-to-reverberant ratio as a function of RT_{60} on %*Alcons*.

The %*Alcons* equations work well with single point or center cluster systems or even split clusters, however, with distributed systems (especially high-density ceiling systems for example) determining the $(n + 1)$ factor becomes extremely difficult, as it is difficult to apportion what percentage of the radiation from adjacent or semiadjacent speakers is actually contributing to the direct field and early fields and which is contributing to the reverberant.

To a certain extent this is made easier in the more complex or long form version as a straight apportionment factor can be applied, though some considerable skill in doing this is required. Because the %*Alcons* equations do not effectively account for the early or late reflected energy, their accuracy needs to be treated with some caution. Furthermore, the method and equations are based on statistical acoustics, which at low reverberation times (e.g., <1.5 s) in itself becomes less accurate.

36.7.2 Intelligibility and Reverberation Time

Although, as we have seen, there is a lot more to intelligibility than reverberation alone, knowing the reverberation time of a space is a good starting point for a system design and immediately allows the potential difficulty of the task to be quantified. Some general rules of thumb can be applied in this context as seen in Table 36-2.

Table 36-1. Effect of Reverberation Time

RT_{60}	**Results**
< 1 s	Excellent intelligibility should be obtained.
1.0–1.2 s	Excellent to good intelligibility should be achieved.
1.2–1.5 s	Good intelligibility should be achieved though loudspeaker type and location become important.
>1.5 s	Careful design required (loudspeaker selection and spacing).
1.7 s	Limit for good intelligibility in large spaces (distributed systems)—e.g., shopping malls, airport terminals.
>1.7 s	Directional loudspeaker required (churches, multipurpose auditoriums, and highly reflective spaces).
>2 s	Very careful design required. High-quality directional loudspeaker required. Intelligibility may have limitations (Concert halls, churches, treated sports halls/arenas.)
>2.5 s	Intelligibility will have limitations. Highly directional loudspeaker required. Large (stone built) churches, sports halls, arenas, atriums, enclosed railway stations, and transportation terminals.
>4 s	Very large churches, cathedrals, mosques, large and untreated atria, aircraft hangars, untreated enclosed ice sports arenas/stadiums. Highly directional loudspeakers required and located as close to the listener as possible.

When designing or setting up systems for use in reverberant and reflective environments, the main rule to follow is, "Aim the loudspeakers at the listeners and keep as much sound as possible off the walls and ceiling." This automatically partially maximizes the direct-to-reverberant ratio, though in practice it may not be quite so simple. The introduction of active and phased line arrays has had a huge impact on the intelligibility that now can be achieved in reverberant and highly reverberant spaces. Arrays of up to 5 m (~16 ft) are readily available and can produce remarkable intelligibility at distances of over 20–30 m even in 10 s plus reverberation time environments. The use of music line arrays has also led to a significant improvement in music/vocal clarity in arenas and concert halls. Whereas the intelligibility form a point or low Q source effectively reduces as square of the distance, this is not the case for a well-designed/installed line array. An example of this is Fig. 36-23 where it can readily be seen that the intelligibility (as measured using the Speech Transmission Index—STI) remains virtually constant over a distance of 30 m in a highly reverberant cathedral (RT = 4 s).

36.8 Some Further Effects of Echoes and Late Reflections

As already noted, speech signals arriving within 35 ms of the direct sound generally integrate with the direct sound and aid intelligibility. In most sound system applications

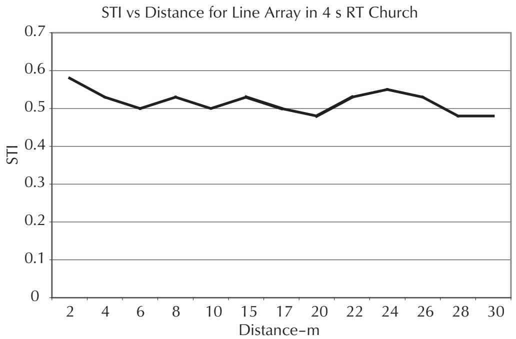


Figure 36-23. Intelligibility versus distance for a four meter line array in a reverberant church.

and particularly in distributed loudspeaker systems, a considerable number of early reflections and sound arrivals will occur at a given listening position. These can provide a useful bridging effect (sequential masking) which can extend the useful arrival time to perhaps 50 ms. The way in which single or discrete reflections affect intelligibility has been studied by a number of researchers—perhaps the best known being Haas.

Haas found that under certain conditions, delayed sounds (reflections) arriving after an initial direct sound could in fact be louder than the direct sound without affecting the apparent localization of the source. This is often termed the *Haas effect.* Haas also found that later arriving sounds may or may not be perceived as echoes depending on their delay time and relative level. These findings are of significant importance to sound system design and enable, for example, delayed infill loudspeakers to be used to aid intelligibility in many applications ranging from balcony infills in auditoria and pew back systems in churches to large venue rear fill loudspeakers. If the acoustic conditions allow, then improved intelligibility and sound clarity can be achieved without loss of localization.

Fig. 36-24 presents a set of echo disturbance curves produced by Haas and shows the sensitivity to disturbance by echoes or secondary sounds at various levels and delay times.

Fig. 36-25, after Meyer and Shodder, shows a curve of echo perception for various delay times and levels (dotted curve) and indicates that delayed sounds become readily discernible at delays in excess of 35 ms (e.g., at 50 ms delay), a single reflection or secondary signal has to be more than 10 dB lower before it becomes imperceptible and has to be more than 20 dB lower at 100 ms. The solid curve in Fig. 36-25 shows when a delayed sound will be perceived as a separate sound source and ceases to be integrated with the direct sound.

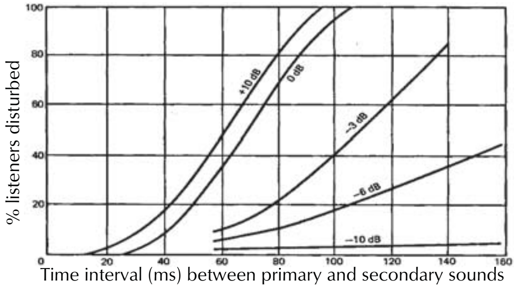


Figure 36-24. Echo disturbance as a function of delay and level (after Haas).

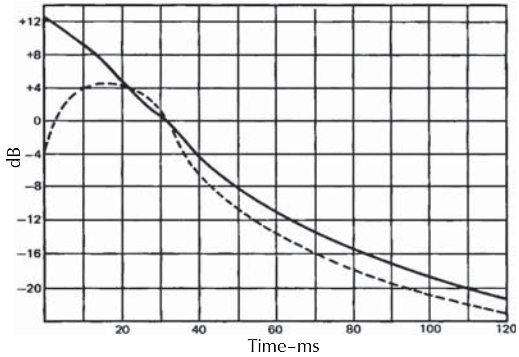


Figure 36-25. Echo perception as a function of delay time and level (after Meyer and Shodder).

Although potentially annoying, echoes may not degrade intelligibility as much as is generally thought. Fig. 36-26, based on work by Peutz, shows the reduction in %*Alcons* caused by discrete sound arrivals or echoes.

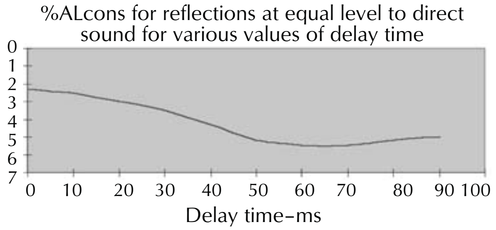


Figure 36-26. Effect of echoes on %*Alcons* (after Peutz).

The curve starts at just under 2% as this was the residual loss due to the particular talker and listener group taking part in the experiment. As the figure shows, the single reflections typically only caused an additional loss of around 2–3%.

However, typically more complex systems operating in reverberant spaces can often give rise to the creation of groups of late reflections which, anecdotally at least, would appear to be rather more detrimental. Fig. 36-27 shows the ETC measured on the stage of a 1000 seat concert hall auditorium. A small group of prominent late reflections is clearly visible.

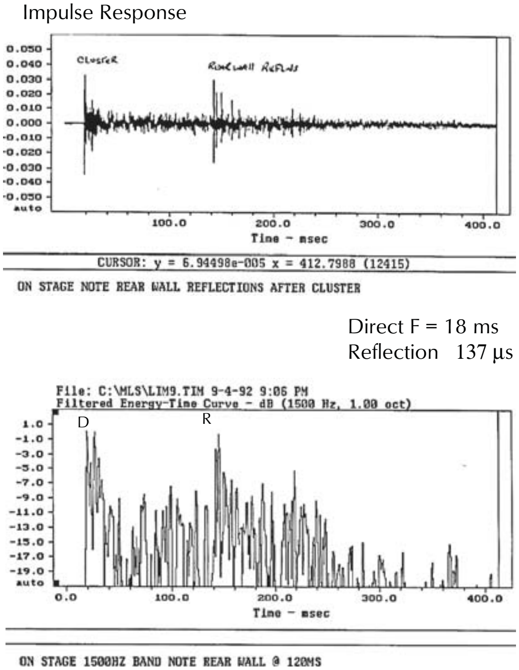
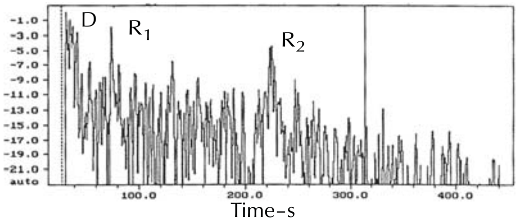


Figure 36-27. Impulse response and ETC of late reflection from auditorium rear wall to stage.

The reflections arrive some 120 ms after the direct sound and are less than 0.5 dB lower and would therefore be expected to be a significant problem. The cause was sound from the center cluster rebounding off the acoustically untreated rear wall and returning to the stage. This was not only clearly audible but also extremely annoying to anyone using the system when speaking from the stage—the coverage and intelligibility throughout the audience area, however, were extremely good. The problem, although clearly caused by the sound system, was in fact, not the fault of the system, but rather the lack of appropriate acoustic treatment on the rear wall. Sound from the cluster had to strike the wall in order to cover the rear rows of seating. Although this was released at the design stage and appropriate treatment arranged, in the event this was not installed and an extremely annoying echo resulted. (Later installation of the specified treatment solved the problem, which shows how important it is to properly integrate systems and acoustics.)

Another interesting problem found in the same auditorium during initial setting up of the system is shown in Fig. 36-28. Again a group of late reflections is clearly visible. A strong reflection occurred 42 ms after the direct sound just 1.9 dB down and the later group arrived 191 ms after the direct and 4.5 dB down.



Side balcony 1500 Hz note reflections @70 ms
R_1 = 42 ms –1.7 dB R_2 = 42 ms –4.5 dB

Figure 36-28. ETC showing late reflections in auditorium causing blurring of sound and loss of intelligibility.

Perhaps surprisingly, the effect of these reflections was not to create a distinct echo but rather to cause a general loss of intelligibility and blurring of the sound. In other nearby seats, the intelligibility was good and measured 0.70 STI but in the seats where the intelligibility was poor the STI was 0.53. Although significantly lower than 0.7, a value of 0.53 would still appear to be too high in relation to the subjective impression obtained. However, Houtgast and Steeneken specifically warn against the use of STI for assessing situations with obvious echoes or strong reflections. Identifying the problem however, would not have been possible without the ability to see the ETC.

36.9 Uniformity of Coverage

It is essential when designing systems to work in noisy and/or reverberant spaces to insure that the direct sound level is as uniform as practical. For example, while a 6 dB variation (±3 dB) may be acceptable under good acoustic conditions, such a variation in a reverberant space can lead to intelligibility variations of 20–40%. A 40% degradation of clarity under such conditions is usually unacceptable. For the case of noise alone, the variation would be at least a 20% reduction in potential intelligibility—though this will be dependent upon the spectrum of the noise. The off-axis performance of a selected loudspeaker is therefore of critical importance—a smooth and well-controlled response being a highly desirable feature.

Where listeners are free to move about—e.g., on a concourse or in a shopping mall—it may be possible to have a greater variation in coverage and hence intelligibility. However, with a seated audience or spectators in an enclosed space, it is essential to minimize seat to seat variations. In critical applications, variations in coverage may need to be held within 3 dB in the 2 kHz and 4 kHz octave bands. This is a stringent and often costly requirement. To put this into perspective consider the following example: assume a given space has an RT_{60} of 2.5 s. Calculation shows that on-axis to the loudspeaker at a given distance gives a value of 10%*Alcons*—an acceptable value. However going off-axis or to a position where the direct sound reduces by just 3 dB will result in a predicted %*Alcons* of 20%—an unacceptable value, see Fig. 36-22. This shows that it is vital to remember off-axis positions as well as the on-axis ones when carrying intelligibility predictions and system designs. Particularly when it is considered that in many applications, the potential intelligibility will be further degraded by the presence of background noise— even when it is not the primary factor.

36.10 Computer Modeling and Intelligibility Prediction

Computer modeling and the current state of the art are discussed in depth in Chapters 9 and 35 and so will only be briefly mentioned here. The ability to accurately predict the direct and reverberant sound fields and compute the complex reflection sequences that occur at any given point are truly remarkable advances in sound system design. As we have seen, calculation of intelligibility from the statistical sound fields alone is not sufficiently accurate for today's needs—particularly with respect to distributed sound systems. The computation of the reflection sequence and hence the impulse response at a point allows far more complex analyses to be carried out including predictions of the early-to-late sound field ratios and the direct calculation of STI. (It should be noted that some of the current simpler programs and many of the earlier prediction programs, although purportedly providing a prediction of STI, in fact base this on a statistical %*Alcons* calculation and convert the resulting value to Rasti. The accuracy of the result value is therefore highly questionable.)

Some program, however, are capable of highly accurate prediction, particularly as the precision of the loudspeaker data increases to ⅓-octave bandwidths and 10 degrees or better angular resolution. Also as the computing power continually increases, greater reflection sequence lengths and orders can be more practically accommodated and hence more accurate reflection field data can be calculated. The main restriction currently is not the mathematical accuracy of the model itself, but the time and effort required to build it in the first place. For many schemes this is simply not economically viable so some form of simple prediction routine, to at least insure that the proposed system will achieve roughly the right order of magnitude of intelligibility, is still required.

36.11 Equalization

It is surprising how many sound systems are still installed either with no or totally inadequate equalization facilities. Yet the major variations in frequency response (both perceived and measured) that systems exhibit when normally installed can have significant effect on the resultant intelligibility and clarity. Equally many systems after they have been equalized often sound worse than they did before.

This is primarily due to a lack of understanding on behalf of the person carrying out the task. There would appear to have been very little research carried out on the effects of equalization on intelligibility. The author has noted improvements of up to 15–20% on some systems, but otherwise the improvements that can be gained are not well publicized.

There are probably about eight main causes of the frequency response anomalies generally observed prior to equalizing a sound system. Assuming that the loudspeaker(s) has a reasonably flat and well-controlled response to begin with these are:

1. Local boundary interactions, Fig. 36-7.
2. Mutual coupling or interference between loudspeakers.
3. Missynchronization of units in a cluster.
4. Incorrectly acoustically loaded loudspeaker, (e.g., a ceiling loudspeaker in too small a back box and/or a coupled cavity).
5. Irregular (poorly balanced) sound power characteristic interacting with reverberation and reflection characteristics of the space.
6. Inadequate coverage, resulting in dominant reverberant sound off-axis.
7. Excitation of dominant room modes (Eigen tones). (These may not appear as large irregularities in the frequency response but subjectively can be very audible and intrusive.)

8. Comparison for high-frequency losses caused by long cable runs or excess atmospheric absorption.

To these may be added abnormal or deficient room acoustics particularly if exhibiting strong reflections or focusing.

Equalization is a thorny subject, with many different views being expressed as to how it should be carried out and what it can and cannot achieve. Suffice it to say that equalization can make a significant improvement to both the intelligibility and clarity of a sound system.

In some cases the improvements are dramatic—particularly when considering not so much the intelligibility per se but associated factors such as ease of listening and fatigue. The essential point is that there is no one universal curve or equalization technique that suits all systems all of the time.

Two examples of this are given below. Fig. 36-29 shows the curves before and after equalization of a distributed loudspeaker system in a highly reverberant church. The anechoic response of the loudspeakers in question is reasonably flat and well extended at high frequencies. Because the measurement (listening) position is beyond the critical distance, the reverberant field dominates and it is the total acoustic power radiated into the space that determines the overall response.

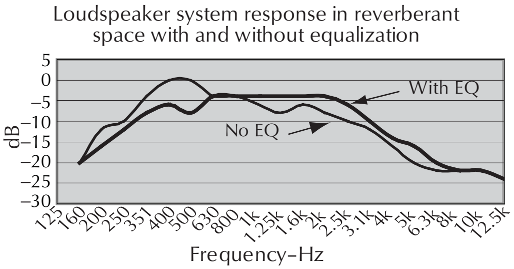


Figure 36-29. Frequency response of a sound system in reverberant church before and after equalization.

The power response of the loudspeaker in question is not flat but falls off with increasing frequency. (This is the normal trend for cone-based devices but some exhibit sharper roll-offs than others.) This, coupled with the longer reverberation time at lower frequencies due to the heavy stone construction of the building, results in an overemphasis at low and lower midfrequencies. The peak at around 400 Hz is due to a combination of power response, mutual coupling of loudspeakers, and boundary interaction effects. The resultant response causes considerable loss of potential intelligibility as high-frequency consonants are lost. Equalizing the system as shown by the solid curve improved the clarity and intelligibility significantly resulting in an improvement of some 15%.

Fig. 36-30 shows a widely quoted equalization curve for speech systems. This has been found to work well for distributed systems in reverberant spaces, but it is only a guideline and should not be regarded too rigorously. Loudspeakers that have a better balanced power response that more closely follows the on-axis frequency response will exhibit less high-frequency roll-off and will generally allow a more extended high-frequency equalization curve.

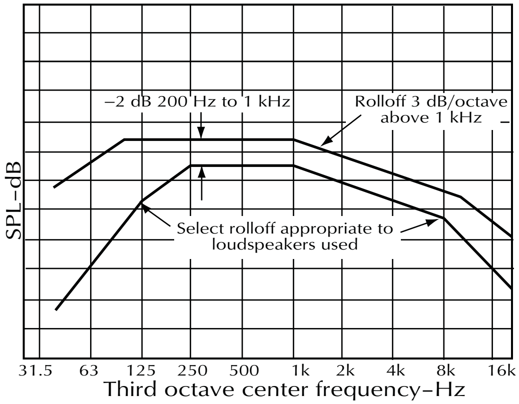


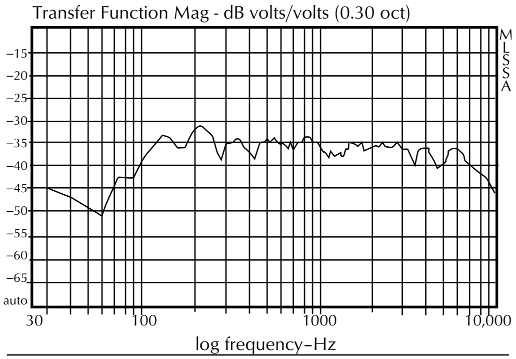
Figure 36-30. Typical response guideline curve for speech reinforcement systems.

An example of this is shown in Fig. 36-31. This is the response of a distributed loudspeaker system employing two-way enclosures in a reflective but well-controlled acoustic environment. In this case, rolling off the high-frequency response would be wholly inappropriate and would degrade the clarity of the system.

Adding bass to a sound system may make it sound impressive but will do nothing for the clarity and intelligibility. Indeed, in general, such an approach will actually reduce the intelligibility and clarity particularly in reverberant spaces. Where music as well as speech need to be played through a system, different paths with different equalization settings should be employed so that the different requirements of each signal can be optimized.

36.12 Talker Articulation and Rate of Delivery

Whereas the sound system designer has some control or at least influence over many of the physical parameters that affect the potential intelligibility of a sound system, an area where no such control exists is that of the person using the microphone. Some talkers naturally articulate



LATERAL LOUDSPEAKER SYSTEM WITH POSITIVE D/A (AFTER EQ)

Figure 36-31. Frequency response curve of distributed two-way loudspeaker system in reflective but well-controlled acoustic space.

better than others and so the resultant broadcast announcements are also inherently clearer.

However, it must not be forgotten that even good talkers cause some loss of potential intelligibility. Peutz, for example, found that good talkers produced 2–3% additional *Alcons* loss over and above that caused by the system and local environment. Poor talkers can produce additional losses of up to 12.5%. It is therefore important to design in some element of safety margin into a sound system in order to compensate for such potential losses.

The rate at which a person speaks over a sound system is also an important factor—particularly in reverberant spaces. Considerable improvement in intelligibility can be achieved by making announcements at a slightly slower than normal rate in acoustically difficult environments such as large churches, empty arenas, gymnasiums, or other untreated venues.

Training announcers or users on how to use the system and how to speak into a microphone can make a significant improvement. The need for proper training can not be overstated but is frequently an area that is often ignored. Prerecorded messages loaded into high-quality, wide bandwidth digital stores can overcome certain aspects of the problem.

For highly reverberant spaces, the speech rate needs to be slowed down from the normal rate of speaking—e.g., from around five syllables per second down to about three syllables per second. This can be very difficult to do under normal operating conditions but carefully rehearsed, slower recordings can be very effective. Equally, the author has found that feeding back a slightly delayed or reverberated signal of the person speaking (e.g., via headphones or an earpiece) can be a very effective way of slowing down the rate of speech—though this has to be carefully controlled and set up, as too much delay can become off putting and counterproductive.

Research has shown that intelligibility is improved when the lips of the talker can be seen. At low levels of intelligibility (e.g., 0.3 to 0.4 AI [Articulation Index]) visual contact can produce improvements of up to 50%. Even with reasonably good intelligibility (e.g., 0.7 to 0.8 AI) improvements of up to 10% have been observed. This suggests that paging and emergency voice alarm systems may have a more difficult task than speech reinforcement systems where additional visual cues are generally also present.

36.13 Summary of Intelligibility Optimization Techniques

The following tips should hopefully prove useful in optimizing sound system intelligibility or act as a catalyst for other ideas and design strategies. Although some would appear very basic, it is remarkable how many systems could be improved with just a minor adjustment or simple redesign.

- Aim the loudspeakers at the listeners and keep as much sound off the walls and ceiling—particularly in reverberant spaces or where long path echoes can be created.
- Provide a direct line of sight between the loudspeaker and listener.
- Minimize the distance between the loudspeaker(s) and listener.
- Insure adequate system bandwidth, extending from a minimum of 250 Hz to 6 kHz and preferably >8–10 kHz.
- Avoid frequency response anomalies and correct unavoidable peaks with appropriate equalization.
- Try to avoid mounting loudspeakers in corners.
- Avoid long path delays (>45 ms). Use electronic signal delays to overcome such problems where loudspeaker spacing >20 ft/6 m (30 ft/9 m max).
- Use directional loudspeakers in reverberant spaces to optimize potential D/R ratios. (Use models exhibiting smoothly controlled and reasonably flat or a gently sloping power response if possible.)
- Minimize direct field coverage variations. Remember that variations of as little as 3 dB can be detrimental in highly reverberant spaces.
- Insure speech SNR is at least 6 dBA and preferably >10 dBA.
- Use automatic noise level sensing and gain adjustment to optimize SNR where background noise is variable.

- Provide a quiet area or refuge for the announcement microphone or use a good quality and effective noise canceling microphone with good frequency response.
- Insure that the microphone user is properly trained and understands the need not to go off mic and to speak clearly and slowly in reverberant environments.
- Repeat important messages.
- In very difficult environments, use simple vocabulary and message formats. Consider use of high-quality specially annunciated prerecorded messages.
- Consider making improvements to the acoustic environment. Do not design the sound system in isolation. Remember, the acoustical environment will impose limitations on the performance of any sound system.

36.14 Intelligibility Criteria and Measurement

A number of intelligibility criteria and rating and assessment methods have already been noted in earlier sections. Here they are treated in a rather more comprehensive overview. However as each technique is quite complex, readers are referred to the bibliography at the end of this chapter to obtain more detailed information.

It is obviously important to be able to specify the desired degree of intelligibility required either for a particular purpose or so that it can be objectively specified for a given project or system. The need then also automatically follows that there has to be a corresponding method of measuring and assessing that a given criterion has been met. Intelligibility measurement and assessment techniques can be divided into two broad categories. These are:

1. Subject based measures—employing a panel of listeners and using a variety of speech-based test materials.
2. Objective acoustic measures of a parameter or parameters that correlate with some aspect of perception.

Subject-based measures include writing down word scores, sentence recognition, modified rhyme tests, and logotom recognition. Objective acoustic measures include broadband and weighted SNR, Articulation Index, Speech Interference Level (SIL and PSIL),direct-to-reverberant measures (including TEF %*Alcons* and C35/C50), and STI. There are also a number of subsets of these latter techniques.

It should not be forgotten that it is not just sound reinforcement or public address systems where the resultant intelligibility may require assessment. Other related audio applications include telephone and intercom systems (telephone/headphone or loudspeaker based) as well as teleconferencing systems and other communication channels—e.g., radio. Hearing assistance systems for the hard of hearing can also be assessed and rated using a number of the techniques described below as can the effectiveness of noise masking systems where conversely a reduction in intelligibility is deliberately sought. Measurements may also need to be made in order to assess the natural intelligibility of a space perhaps so that the potential benefits or need for a speech reinforcement system can be evaluated and objectively rated (e.g., churches, classrooms and lecture rooms/ auditoria, etc.).

Not all of the techniques are applicable to every application. The area of application is therefore noted at the end of each section. The practical limitations of each of the methods described are also briefly discussed.

36.14.1 Subject-Based Measures and Techniques

The fundamental measurement of intelligibility is of course speech itself. A number of techniques have been developed to rate speech intelligibility. The initial work was carried out in the 1920s and 1930s and was associated with telephone and radio communication systems. From this work the effects of noise, SNR, and bandwidth were established and subjective test methods formulated. (Much of this work was carried out at Bell Labs under the direction of Harvey Fletcher.) The sensitivity of the various test methods was also established and it was found that tests involving sentences and simple words were the least sensitive to corruption but often did not provide sufficiently detailed information to enable firm conclusions to be drawn regarding the effects and parameters under study to be definitely made.

The need to insure that all speech sounds were equally included led to the development of phonemically balanced (PB) word lists. Lists with 32, then 250, and finally 1000 words were developed. Tests using syllables (logatoms) were also developed. These latter tests provide the most sensitive measure of speech information loss but are complex and very time consuming and costly in application.

The modified rhyme test (MRT) was developed as a simpler alternative to PB word lists and is suitable for use in the field with only a short training period. (The more sensitive methods can require several hours of training of the subjects before the actual tests can begin.) The various methods and their interrelationships are shown in Fig. 36-12 where the Articulation Index is used as the common reference.

36.14.2 Objective Measures and Techniques

36.14.2.1 Articulation Index

The Articulation Index (AI) was one of the first criteria and assessment methods developed to use acoustic measurements and relate these to potential intelligibility. AI is concerned with rating the effects of noise on intelligibility and was primarily developed for assessing telephone communication channels. Later corrections were added in attempt to take account of room reverberation but these methods are not considered sufficiently accurate for sound system use. AI is a very accurate and useful method of assessing and rating the effects of noise on speech. ANSI Standard S3.5 1969 (subsequently revised in 1988 and 1997) specifies the methods of calculation based on measurements of the spectrum of the interfering noise and desired speech signal. (Either in terms of ⅟₁-octave or ⅓-octave bands.)

The Index ranges from 0 to 1 with 0 representing no intelligibility and 1 representing 100% intelligibility. The Index is still very good for assessing the effects of noise on speech in range of applications where room reverberation effects are negligible—e.g., communications channels or aircraft cabins, etc.

Another important application relates to the assessment of speech privacy in offices and commercial environments. Here a very low AI score is required in order to insure that neighboring speech is not intelligible. This is extremely useful when setting up and adjusting sound masking systems and a speech privacy scale has been developed for this purpose. Unfortunately, few commercial analyzers incorporate the measurement, which would be an extremely simple matter to do if a ⅓-octave real-time spectrum display and data are available. Currently, most users of AI in this application either have to compute the result manually or by a simple spreadsheet procedure.

36.14.2.2 Articulation Loss of Consonants

This method was developed by Peutz during the 1970s and further refined during the 1980s. The original equation is simple to use and is in fact based on a calculation of the D/R ratio, although this is not immediately obvious from the equation. The long form of the equation takes into account both noise and reverberation—but unfortunately does not give exactly similar values to the simpler form—which is regarded by many to be overly optimistic. The original work was based on human talkers and not sound systems. (The original prediction equation was modified by Klein in 1971 to its now familiar form in order to do this.)

During 1986, a series of speech intelligibility tests were run that enabled a correlation to be found between MRT word scores carried out under reverberant conditions and a D/R measurement carried out on the TEF analyzer. For the first time this allowed the widely used predictive and design rating technique to be measured in the field. However the correlation does have a number of limitations which need to be considered when applying the method. The measurement bandwidth used at the time was equivalent to approximately ⅓ octave centered at 2 kHz. Although three very different venues were employed, each with three significantly different loudspeakers and directivities, the correlation and hence method is only valid for a single source sound system. The measurement requires considerable skill on behalf of the operator in setting up the ETC measurement parameters and divisor cursors, so a range of apparently correct answers can be obtained. Nonetheless the measurement does provide a very useful method of assessment and analysis. In 1989 Mapp and Doany proposed a method for extending the technique to distributed and multiple source sound systems by extending the duration of the measurement window out to around 40 ms.

A major limitation of the method is that it only uses the 2 kHz band. For natural speech where there is essentially uniform directivity between different talkers, single band measurements can be acceptably accurate. However, the directivity of patterns of many if not the majority of loudspeakers used in sound systems is far from constant and can vary significantly with frequency—even over relatively narrow frequency ranges. Equally, by only measuring over just one narrow frequency band, no knowledge is obtained regarding the overall response of the system. The accuracy of the measurement correlation can therefore become extremely questionable and any apparent %*Alcons* values extracted must be viewed with caution.

36.14.2.3 Direct-to-Reverberant and Early-to-Late Ratios

Direct-to-reverberant measurements or more accurately direct and early reflected sound energy-to-late reflected and reverberant energy ratios have been used as predictors of potential intelligibility in architectural and auditorium acoustics for many years. A number of split times have been employed as delineators for the direct or direct and early reflected sounds and the late energy. The most common measure is C50, which takes as its ratio the total energy occurring within the first 50 ms to the total sound energy of the impulse response. Other measures include C35, whereby the split time is taken as 35 ms and also sometimes C7 where this early split time effectively produces an almost pure D/R ratio.

A well-defined scale has not been developed, but it is generally recommended that for good intelligibility (in an auditorium or similar relatively large acoustic space) a positive value of C50 is essential and that a value of around +4 dB C50 should be aimed for. (This is equivalent to about 5%*Alcons*.) Measurements are usually made at 1 kHz or may be averaged over a range of frequencies. The method does not take account of background noise and is of limited application with respect to sound systems due to the lack of a defined scale and frequency limitations—although there is no reason why the values obtained at different frequencies could not be combined in some form of weighted basis. (See Lochner and Burger 1964.) Bradley has extended the C50 and C35 concept and introduced U50 and U80 etc. where U stands for useful energy. He also included signal-to-noise ratio effects. While the concept is a useful addition to the palette of speech intelligibility measures, it has not caught on to any extent—but it can be a very useful diagnostic tool and further extends our knowledge and understanding of speech intelligibility.

36.14.2.4 Speech Transmission Index STI, RASTI, and STIPA

The STI technique was also developed in Holland at about the same time as Peutz was developing %*Alcons*. While the %*Alcons* method became popular in the United States, STI became popular and far more widely used in Europe and has been adopted by a number of International and European Standards and codes of practice relating to sound system speech intelligibility performance as well as International Standards relating to aircraft audio performance. It is interesting to note that while %*Alcons* was developed primarily as a predictive technique, STI was developed as a measurement method and is not straightforward to predict! (See later.)

The technique considers the source/room (audio path)/listener as a transmission channel and measures the reduction in modulation depth of a special test signal as it traverses the channel, Figs. 36-32 and 36-33. A unique and very important feature of STI is that it automatically takes account of both reverberation and noise effects when assessing potential intelligibility.

Schroeder later showed that it is also possible to measure the modulation reduction and hence STI via a system's impulse response. Modern signal processing techniques now allow a variety of test signals to be used to obtain the impulse response and hence compute the STI—including speech or music. A number of instruments and software programs are currently available that enable STI to be directly measured. However, care needs to be taken when using some programs to insure that any background or interfering noise is properly accounted for.

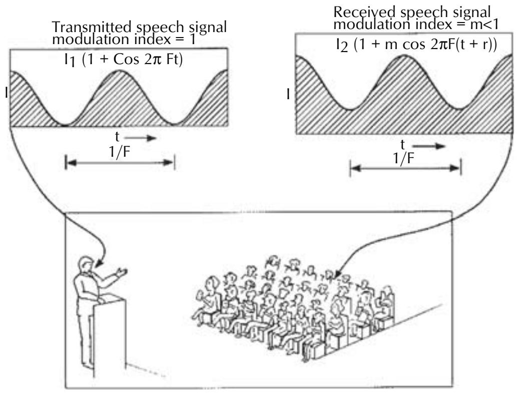


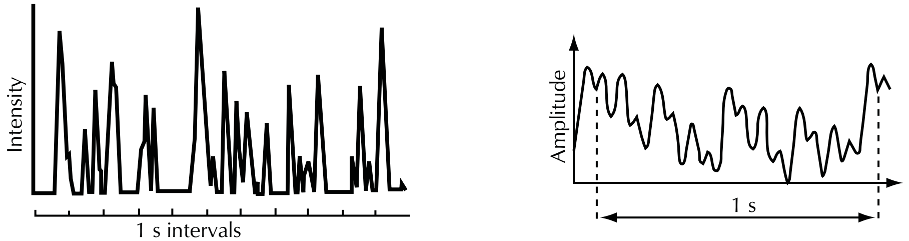
Figure 36-32. Principle of STI and modulation reduction of speech by room reverberation.

The full STI technique is a very elegant analysis method and is based on the amplitude modulations occurring in natural speech, Figs. 36-33 and 36-34. Measurements are made using octave band carrier frequencies of 125 Hz to 8 kHz, thereby covering the majority of the normal speech frequency range. Fourteen individual low-frequency (speechlike) modulations are measured in each band over the range 0.63 to 12.5 Hz.

A total of 98 data points are therefore measured for each STI value (7 octave band carriers each × 14 modulation frequencies). Because the STI method operates over almost the entire speech band it is well suited to assessing sound system performance. The complete STI data matrix is shown in Table 36-2. "X" represents a data value to be provided.

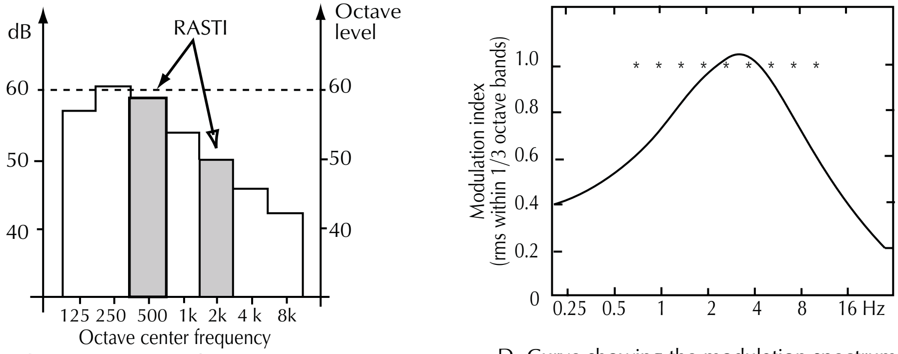
When STI was first developed, the processing power to carry out the above calculations was beyond economic processor technology and so a simpler derivative measure was conceived—RaSTI. RaSTI stands for Rapid Speech Transmission Index (later changed to Room Acoustic Speech Transmission Index when its shortfalls for measuring sound system performance were realized (see Mapp 2002 and 2004). RaSTI uses just nine modulation frequencies spread over two octave band carriers thereby producing an order of magnitude reduction in the processing power required.

The octave band carriers are 500 Hz and 2 kHz, which, although well selected to cover both vowel and consonant ranges, does mean that the system under test



A. An example of the intensity envelope of a segment of human speech.

B. RASTI signal modulation (as applied in the 2 kHz octave).



C. A long-term averaged octave spectrum of normal human speech, at 1 m distance ($L_{eq.,A}$ = 60 dB). The shaded portions indicate the carrier signal used in the RASTI method.

D. Curve showing the modulation spectrum of human speech. The discrete modulation frequencies used in the RASTI method are marked " * ". Four modulation frequencies are applied in the 500 Hz octave and five in the 2 kHz octave as follows:
500 Hz octave: 1 Hz, 2 Hz, 4 Hz, 8 Hz
2 kHz octave: 0.7 Hz, 1.4 Hz, 2.8 Hz, 5.6 Hz, 11.2 Hz.

Figure 36-33. Principle of STI and RASTI showing octave band spectrum and speech modulation frequencies.

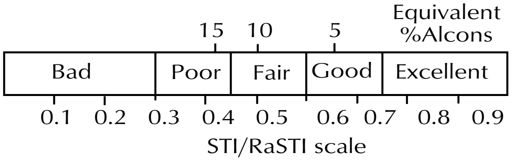


Figure 36-34. STI subjective scale and comparison with %Alcons.

has to be reasonably linear and exhibit a well-extended frequency response. Unfortunately many paging and voice alarm systems do not fulfill these criteria and so can give rise to readings of questionable accuracy. (However, this still takes account of a wider frequency range than the traditional D/R and %*Alcons* methods.) Fig. 36-35 shows a system response simulated by the author and evaluated via RaSTI. Although the majority of the speech spectrum is completely missing, the result was an almost perfect score of 0.99 STI!

Table 36-2. STI Modulation Matrix

Carrier/Modulation Frequency (Hz)	125	250	500	1K	2K	4K	8K
0.63	X	X	X	X	X	X	X
0.80	X	X	X	X	X	X	X
1.0	X	X	X	X	X	X	X
1.25	X	X	X	X	X	X	X
1.6	X	X	X	X	X	X	X
2.0	X	X	X	X	X	X	X
2.5	X	X	X	X	X	X	X
3.15	X	X	X	X	X	X	X
4.0	X	X	X	X	X	X	X
5.0	X	X	X	X	X	X	X
6.3	X	X	X	X	X	X	X
8.0	X	X	X	X	X	X	X
10.0	X	X	X	X	X	X	X
12.5	X	X	X	X	X	X	X

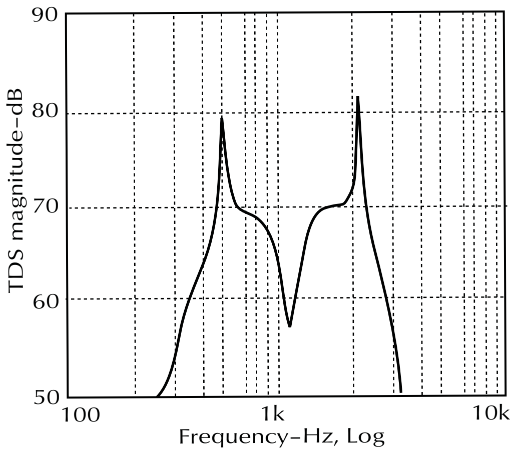


Figure 36-35. Simulated system frequency response curve favoring 500 and 2000 Hz, giving an excellent RASTI value but very poor sound quality.

The first commercially available instrument that could measure STI was the RaSTI meter introduced by Brüel and Kjaer in 1985. Modulated pink noise in the 500 Hz and 2 kHz octaves is generated and transmitted either acoustically or electronically into the system under test. A useful feature of the method is that the resultant signal is very speechlike in nature having a crest factor of around 12 dB, which compares well with normal speech (at around 15–20 dB). The relative levels of the 500 Hz and 2 kHz signals are also automatically transmitted in their correct ratios as compared to natural speech.

This makes setting up the test signal levels comparatively straightforward and enables measurements to be carried out by trained but nonexpert personnel. The introduction and adoption of the RaSTI test method has literally revolutionized the performance of many PA systems ranging from aircraft cabins and flight recording systems to trains, malls, and cathedrals as, for the first time, the intelligibility of such systems could be set and readily verified. As the limitations of RaSTI as a measure of sound system performance became more widely known and understood (e.g., Mapp 2002 and 2004) it became clear that a replacement method would be required. In 2001, STIPA (STI for PA systems) was introduced and became part of IEC 268-16 in 2003. Unlike RaSTI STIPA measures over virtually the complete speech bandwidth range from 125 Hz to 8 kHz. However, a sparse matrix is used to cut down the complexity of the stimulus and associated measurement processing time. Table 36-3 shows the modulation matrix for STIPA.

Table 36-3. STIPA Modulation Matrix

Carrier/Modulation Frequency (Hz)	125	250	500	1K	2K	4K	8K
0.63			X				
0.80						X	
1.0	X	X					
1.25					X		
1.6	?						
2.0				X			
2.5							X
3.15			X				
4.0						X	
5.0	X	X					
6.25					X		
8.0	?						
10.0				X			
12.5							X

The current version of IEC 268-16 (Edition 3, 2003) employs the above modulations. It can be seen that for some reason the 125 Hz and 250 Hz carriers employ the same modulation frequencies—although there is a spare set available (there are no modulations at 1.6 Hz and 8 Hz). However, it is quite likely that a future version of the standard (2010) will correct this apparent anomaly and use the missing modulations (indeed some meters already do this).

Since its introduction, the STIPA technique has rapidly taken off with at least four manufacturers offering handheld portable measurement devices (though some are more accurate than others—see Mapp 2005). In a similar manner to RaSTI, the STIPA signal is speech shaped and so automatically presents the correct signal for assessing a sound system. (STIPA is a protected name and refers to a measurement made with a modulated signal. Although STIPA can be derived from an impulse response, any such measurement must be clearly indicated as being an equivalent STIPA). At the time of writing, the STIPA signal has been relatively loosely defined, but it is understood that the fourth edition of IEC 268-16 will clarify the issue. This should also help insure that the various STIPA meters are fully compatible, so that any 2 meters using the same IEC268-16 Ed 4 test signal should give the same result.

The typical time required to carry out a single STIPA measurement is around 12–15 s. However, as the test signal is based on a pseudorandom signal, there can be some natural variation between readings. For this reason it is recommended that at least three readings be taken to insure that a reliable measurement result is produced. STIPA correlates very closely with STI and overcomes most of the problems associated with RaSTI. However, just as with STI, it is a far from perfect measure and there are a number of limitations that need to be understood.

STIPA is vulnerable to some forms of digital signal processing and in particular CD player errors. STIPA test signals are generally distributed on CDs, and some CD players can introduce significant errors. It is therefore essential to conduct a loop back measurement to insure that a valid signal is being generated. More recently, however, signals have been distributed as wav files on solid state memory cards and can also can be directly downloaded, which helps overcome the problem (see Mapp 2005 for further details). Most hardware implementations include some form of error detection—particularly the detection of the occurrence of impulsive noise during a test. Not all STIPA meters have incorporated the level dependency relationship that exists between speech level and intelligibility—although it is clearly defined within the IEC 268-16 standard. The same is also true of the masking function that the standard requires.

One of the major shortcomings of STI and STIPA is their inability to correctly assess the effect that equalization can have on a sound system. For example, STI measurements made on the system described earlier and whose frequency response is depicted in Fig. 36-27 were exactly the same pre and post equalization—although the word score intelligibility improved significantly. Adapting STI to correctly account for such occurrences is not a simple or straightforward matter and it will be some time yet before we have a measure that can accurately do this.

The STI/RaSTI scale ranges from 0 to 1. Zero represents total unintelligibility while 1 represents perfect sound transmission. Good correlation exists between the STI scale and subject-based word list tests. As with all current objective electroacoustic measurement techniques, STI does not actually measure the intelligibility of speech, but just certain parameters that correlate strongly with intelligibility. It also assumes that the transmission channel is completely linear. For this reason, an STI measurement can be fooled by certain system nonlinearities or time-variant processing. STI is also liable to corruption by the presence of late discrete arrivals (echoes). These, however, can be readily spotted by examination of the modulation reduction matrix.

The basic equation for the STI modulation reduction factor *m*(*f*) is

$$m(f) = \frac{1}{1+\left[\frac{2FT}{13.8}\right]^2} \times \frac{1}{1+10^{\frac{\text{,ÄìS/N}}{10}}} \qquad (36\text{-}5)$$

Unfortunately, this equation cannot be directly solved, making STI prediction a complex procedure requiring detailed computer modeling and analysis of the sound field. An approximate relationship exists between STI (RASTI) and %*Alcons*. Fig. 36-34 shows the two scales while Table 36-4 gives a numerical set of equivalent values.

Table 36-4. RaSTI and %*Alcons* Numerical Set of Equivalent Values

Quality	RASTI	%*Alcons*	Quality	RASTI	%*Alcons*
	0.20	57.7		0.62	6.0
	0.22	51.8		0.64	5.3
	0.24	46.5		0.66	4.8
	0.26	41.7		0.68	4.3
BAD	0.28	37.4		0.70	3.8
	0.30	33.6		0.72	3.4
	0.32	30.1		0.74	3.1
	0.34	27.0	GOOD	0.76	2.8
	0.36	24.2		0.78	2.5
	0.38	21.8		0.80	2.2
	0.40	19.5		0.82	2.0
	0.42	17.5		0.84	1.8
POOR	0.44	15.7		0.86	1.6
	0.46	14.1		0.88	1.4
	0.48	12.7		0.90	1.3
	0.50	11.4		0.92	1.2
	0.52	10.2	EXCELLENT	0.94	1.0
	0.54	9.1		0.96	0.9
	0.56	8.2		0.98	0.8
FAIR	0.58	7.4		1.0	0.0
	0.60	6.6			

The subjective scale adopted for STI (and RaSTI/STIPA) has led to considerable confusion when rating PA and sound systems. (For example a rating of 0.5 STI would normally be rated as good if heard in a highly reverberant or difficult acoustic environment rather than only fair. Also, in practice, there is usually a marked difference in perception between 0.45 and 0.50 STI (and more particularly 0.55 STI)—although they are all rated as fair. In an attempt to overcome the problem and also to add a degree of tolerance to the measure, the author has proposed that a new rating scale be employed for PA/sound systems (Mapp 2007). The proposed scale is shown in Fig. 36-36 and is based on a series of designated bands rather than absolute categories. While the bands will remain fixed, their application can vary so that, for example, an emergency voice announcement system may be required to meet category "G" or above whereas a high-quality system for a theater for a concert hall might be required to meet category "D," or an assistive hearing system might be required to meet category "B" or above, etc. (Table 36-5). It is anticipated that the new scale will be adopted by IEC 268-16 and form part of the fourth edition of the standard.

Table 36-5. Possible Rating Scheme for Sound Systems

Category	Typical Use	Comment
A		
B	Theaters, speech auditoria, HOH systems	High speech intelligibility,
C	Theaters, speech auditoria, teleconferencing	High speech intelligibility,
D	Lecture theaters, classrooms, concert halls, modern churches	Good speech intelligibility,
E	Concert halls, modern churches	High-quality PA systems,
F	Shopping malls, public buildings offices, VA systems	Good quality PA systems
G	Shopping malls, public buildings offices, VA systems	Target requirement for VA/PA
H	VA & PA systems in difficult acoustic environments	Lower target for VA/ PA
I	VA & PA systems in difficult spaces	
J	Not suitable for PA systems	
U	Not suitable for PA systems	

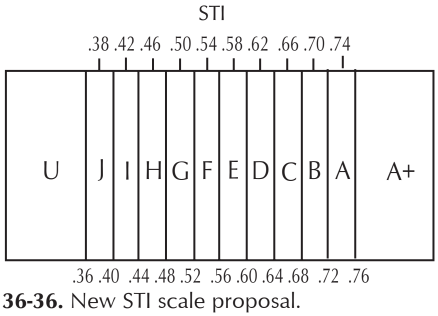


Figure 36-36. New STI scale proposal.

Whereas STI does incorporate a degree of diagnostic ability—e.g., it can readily be determined if the speech intelligibility is primarily being reduced by noise or reverberation and the presence of late reflections—a visual display of the ETC or impulse response is invaluable when actually determining appropriate remedial measures and identifying the underlying cause or offending reflective surfaces. A combination of techniques is therefore employed by the author when tackling such problems including the use of directional and polar ETC measurements.

36.14.2.5 SII Speech Intelligibility Index

This relatively new index *(*ANSI S3.5 1997*)* is closely related to the Articulation Index (AI) but also makes use of some STI concepts. SII calculates the effective signal-to-noise ratio for a number of frequency bands related to the speech communication bandwidth. Several procedures with different frequency domain resolutions are available. These include conventional ⅓-octave and 1⁄1-octave as well as a twenty one band critical bandwidth (ERB) analysis. An analysis based on seventeen equally contributing bands is also incorporated. The method would appear to be more suitable for direct communication channels rather than for sound reinforcement and speech announcement systems, but in situations where reverberation has little or no effect, the method would be applicable. It should also be useful for evaluating and quantifying the effectiveness of speech masking systems.

36.14.3 The Future for Speech Intelligibility Measurements

As can be seen from the foegoing discussions, we are still a long way from truly measuring speech intelligibility itself. Currently all we can do is to measure a number of physical parameters than correlate under certain conditions to intelligibility. An order of magnitude improvement is required for these to become less anomalous and fallible. The power of the modern PC should allow more perceptually based measurements to be made, as is already happening with telephone networks. However, it must not be forgotten that what is needed in the field is a simple to operate system that does not require highly trained staff to operate, as one thing is for certain: the need to measure and prove that intelligibility criteria have been met is going to rapidly expand over the next few years—indeed the introduction of STIPA is already hastening the process. The range of applications where such testing will need to be performed is also going to rapidly expand and will encompass almost all forms of public transport as well as all forms of voice-based life safety systems. The more traditional testing of churches, auditoria, classrooms, stadiums, and transportation terminals is also set to rapidly expand. As DSP technology continues to grow and as our understanding of psychoacoustics and speech continues to develop, the ability to manipulate speech signals to provide greater intelligibility will increase. Measuring such processes will be a further and interesting challenge.

Particular areas that are likely to see progress over the next few years are the development and use of binaural intelligibility measurements—probably using STI as their basis. The author has also tried using STI and STIPA to assess speech privacy and the effectiveness of speech masking systems. While potentially a promising technique, there are still many obstacles to overcome before it can become a viable technique—not least of which requires considerable research to be carried out between speech intelligibility and STI at the lower end of the STI scale (Mapp 2007). The measurement and intelligibility assessment of assistive hearing systems is also currently under investigation (Mapp 2008) and is showing considerable promise. It is anticipated that a series of new criteria and measurement techniques will be developed specifically for this specialized but increasingly important field. The use of real speech and other conventional PA signals is also under research and should pave the way for less invasive measurement techniques.

Acknowledgments

The author gratefully acknowledges the assistance of JBL Professional for allowing a number of graphs and figures originally created for their Technical Note “Speech Intelligibility” (TN, vol. 1, no. 26) to be reproduced in this chapter. The author would also like to thank Don and Carolyn Davis for holding the first Syn-Aud-Con Speech Intelligibility Workshop in 1986, which has acted as an inspiration to further work on this subject ever since.

Bibliography

W. Ahnert and F. Steffens, *Sound Reinforcement Engineering*, London, Spon: 1999.

J. Borwick (ed.), *Loudspeaker and Headphone Handbook*, London, Butterworths: 1988, and 3rd Edition 2000.

J. Bradley “Relationships Among Measures of Speech Intelligibility in Rooms,” *JAES,* vol. 48, no. 5, 1998.

Don and Carolyn Davis, *Sound System Engineering,* IN, Howard Sams: 1987.

J. Eargle, *Sound System Design Reference Manual,* JBL Professional: 1999.

J. Eargle, *Electroacoustical Reference Data,* Van Nostrand Reinhold, 1994.

B. J. More, *An Introduction to the Psychology of Hearing,* Academic Press, 1988.

B. S. Atal, M. R. Schroeder, and G. M. Sessler, “Subjective Reverberation Time and Its Relation to Sound Decay,” *5th International Congress on Acoustics*, Liege, 1965.

M. Barron, “The Subjective Effects of First Reflections in Concert Halls—The Need for Lateral Reflections,” *Journal of Sound Vibration*, Vol. 15, no. 4, 1971.

M. F. E. Barron, “The Effects of Early Reflections on Subjective Acoustical Quality in Concert Halls,” Thesis, Southampton University, 1974.

F. A. Bilsen and I. Kievits, “The Minimum Integration Time of the Auditory System,” *AES preprint #2746*, 1989 Convention.

Jens Blauert, *Spatial Hearing, the Psychoacoustics of Human Sound Localization*, Cambridge, MA: The MIT Press, 1983.

E. Corliss, “The Ear as a Mechanism of Communication,” *J. Audio Eng. Soc.*, Vol 38, No. 9, 1990.

Don and Carolyn Davis, “Application of Speech Intelligibility to Sound Reinforcement,” *J. Audio Eng. Soc.*, Vol 37, no. 12, December 1989.

Don and Carolyn Davis, *Sound System Engineering*, 2nd ed. Carmel, IN: SAMS, a Division of Macmillan Computer Publishing, 1986.

Carolyn Davis, “Measurement of %Alcons,” *J. Audio Eng. Soc.*, Vol. 34, no. 11, November, 1986, pp. 905–909.

Don and Carolyn Davis, “Speech Intelligibility Workshop,” *Syn-Aud-Con Tech Topic,* Vol. 14, no. 1, Fall 1986, and vol. 14, no. 8, Summer 1987.

H. Haas, ‘The Influence of a Single Echo on the Audibility of Speech,” *Acustica*, Vol. 1, no. 49, 1951, reprinted *JAES*, 20, 1972.

H. G. Latham, “The Signal-to-Noise Ratio for Speech Intelligibility—An Auditorium Design Index,” *Applied Acoustics*, vol. 12, July 1979.

J. P. A. Lochner and J. F. Burger, “The Intelligibility of Reinforced Speech,” *Acustica* 9, 1959.

J. P. A. Lochner and J. F. Burger, “The Influence of Reflections on Auditorium Acoustics,” *Journal of Sound Vibration,* vol. 4, October 1964.

W. Klein, “Articulation Loss of Consonants as a Basis for Design and Judgment of Sound Reinforcement Systems,” *J. Audio Eng. Soc.*, Vol. 19, no. 11, pp. 920–922, December 1971.

F. Li and T. Cox, “Extraction of STI from Speech Signals Using Artifical Neutral Networks,” *AES110th Convention*, Amsterdam, 2001.

J. P. A. Lochner and J. F. Burger, (*Acustica,* 8,1,)1958.

P. Mapp and P. Doany, “Speech Intelligibility Analysis and Measurement for a Distributed Sound Sysem in a Reverberant Environment,” *AES 87th Convention*, New York, 1989.

P. Mapp, “How Loud is Loud Enough?, *Sound & Video Contractor*, Vol. 14, no. 1, January 1996.

P. Mapp, “Practical Limitations of Objective Speech Intelligibility Measurements,” *AES 102nd Convention*, Munich, 1997.

P. Mapp, “Some Effects of Equalisation and Spectral Distortion on Speech Intelligibility,” *Proc. IOA,* vol. 19, pt. 6, 1997.

P. Mapp, “Objective Speech Intelligibility Testing of Sound Systems,” *Proc. IOA* , Vol. 21, pt. 5, 1999.

P. Mapp, “Reaching the Audience,” *Sound & Video Contractor*, Vol. 17, no. 11, October 1999.

P. Mapp, “Sound Out of Doors,” *Sound & Video Contractor*, Vol. 18, no.1, January 2000.

P. Mapp, “EQ article,” *Sound & Video Contractor*, Vol. 18, no.14, Dec. 2000.

P. Mapp, “Investigations into Problems in Auditorium Acoustics,” *Proc. IOA*, vol. 17, pt. 1, 1995.

P. Mapp, “From Loudspeaker to Ear—Measurement and Reality,” *AES UK Conference*, 1977.

P. Mapp, “Sound Power: The Forgotten Loudspeaker and Sound System Parameter,” *AES 110th Convention*, Amsterdam, 2001.

P. Mapp, “The loudspeaker-Room Interface,” *Hi-fi News & Record Review,* vol. 20, no. 8, 1981.

P. Mapp, “Is STIPA a Robust Measure of Speech Intelligibility Performance,” *AES 118th Convention*, Barcelona, 2005.

P. Mapp, “Relationships between Speech Intelligibility Measures for Sound Systems,” *AES 112th Convention*, Munich, 2002.

P. Mapp, “Limitations of Current Sound System Intelligibility Verification Techniques,” *AES 113th Convention*, Los Angeles, 2002.

P. Mapp, “Systematic and Common Errors in Sound System STI and Intelligibility Measurements,” *AES 117th Convention*, San Francisco, 2004.

P. Mapp, “ How Accurate Are Acoustic Models & Measurements with Regard to Speech Intelligibility?,” *Proc. IOA,* vol. 29, Pt.7, 2007.

P. Mapp, “Assessing the Potential Intelligibility of Assistive Audio Systems for the Hard of Hearing and Other Users,” *AES 124th Convention*, Amsterdam, 2008.

E. Meyer and C. R. Schodder, Gottinger Nachrichten Math. Phys. Kl., 11a, 1962.

V. M. A. Peutz, “Speech Information and Speech Intelligibility,” *AES 85th Convention,* 1988.

V. M. A. Peutz, “Articulation Loss of Consonants as a Criterion for Speech Transmission in a Room,” *JAES*, Vol. 19, no. 11, 1971.

V. M. A. Peutz, “Speech Intelligibility,” *AES 75th Convention,* Paris, 1984.

H. J. Steeneken and T. Houtgast, “Some Applications of Speech Transmission Index (STI) in Auditoria,” *Acustica,* 51, pp. 229–234, 1982.

H. J. Steeneken and T. Houtgast, “Rasti: A Tool for Evaluating Auditoria,” *Brüel & Kjaer Technical Review,* no. 3, 1985.

T. Houtgast, H. J. M. Steeneken, and R. Plomp, “Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function,” *Acustica*, Vol. 46, September, 1980.

H. Wallach et al., *Am. J. Psychology*, Vol. 62, no. 315, 1949.

B. Yegnanarayana and B. S. Ramakrisina, “Intelligibility of Speech under Non-Exponential Decay Conditions,” *Journal of the Acoustical Society of America*, Vol. 58, 1975.

ANSI Standard S3.5- 1969 / 1988 / 1997 Methods for the Calculation of the Articulation Index.

IEC 289/BS 60286–16 Objective Rating of Speech Intelligibility by Speech Transmission Index, 1998.

This page intentionally left blank

Chapter 37

Personal Monitor Systems

by Gino Sigismondi

37.1 Background 1415
37.2 Personal Monitor System Advantages 1416
37.3 Sound Quality 1416
  37.3.1 Adequate Volume 1416
  37.3.2 Gain-Before-Feedback 1417
  37.3.3 Hearing Conservation 1418
  37.3.4 Portability 1418
  37.3.5 Mobility 1418
  37.3.6 Personal Control 1419
37.4 Choosing a System 1420
37.5 Configuring a Personal Monitor System 1420
  37.5.1 How Many Mixes Are Required? 1420
  37.5.2 Stereo or Mono? 1421
  37.5.3 How Many Mixes Are Available from the Mixing Console? 1422
  37.5.4 How Many Components Are Needed? 1422
37.6 Earphones 1422
  37.6.1 Earphone Options 1422
  37.6.2 Earphone Transducers 1424
  37.6.3 The Occluded Ear 1425
  37.6.4 Ambient Earphones 1425
37.7 Applications for Personal Monitors 1425
  37.7.1 Rehearsals 1425
    37.7.1.1 Performance, Club/Corporate/Wedding Bands—No Monitor Mixer 1426
    37.7.1.2 Club Level Bands—With Monitor Console 1427
    37.7.1.3 Professional Touring System 1427
  37.7.2 Mixing for Personal Monitors 1427
  37.7.3 Stereo Wireless Transmission 1428
  37.7.4 Personal Monitors for Houses of Worship and Sound Contractors 1429
37.8 Expanding the Personal Monitor System 1430
  37.8.1 Personal Monitor Mixers 1430
  37.8.2 Distributed Mixing 1430
  37.8.3 Supplementary Equipment 1430
    37.8.3.1 Drum Throne Shakers 1430
    37.8.3.2 Ambient Microphones 1431
    37.8.3.3 Effects Processing 1431
  37.8.4 Latency and Personal Monitoring 1431
  37.8.5 Safe Listening with Personal Monitors 1432
    37.8.5.1 Use an Isolating Earphone 1433
    37.8.5.2 Use Both Earphones! 1433

37.8.5.3 Keep the Limiter On 1434
37.8.5.4 Pay Attention to What Your Ears Are Telling You 1434
37.8.5.5 Have Your Hearing Checked Regularly 1434
For more information on hearing health for musicians, contact one of the following organizations.. 1435
References 1435

37.1 Background

The emergence of modern sound reinforcement systems for music in the 1960s brought with it the need for performers to be able to better hear themselves onstage. Prior to the days of arena concerts and stacks of Marshall™ amplifiers, it wasn't that difficult for performers to hear vocals through the main PA loudspeakers. Most concerts were held in smaller venues, with a few notable exceptions. When the Beatles played Shea Stadium in 1964, the only PA was for voice; guitars were only as loud as the guitar amplifiers. Of course, the crowd noise was so loud even the audience couldn't hear what was going on, let alone the band! As rock and roll shows continued to get bigger and louder, it became increasingly difficult for performers to hear what they were doing. The obvious solution was to turn some of the loudspeakers around so they faced the band. A further refinement came in the form of wedge-shaped speakers that could be placed on the floor, facing up at the band, finally giving singers the ability to hear themselves at a decent volume, Fig. 37-1. With the size of stages increasing, it became difficult to hear everything, not just the vocals. Drums could be on risers 15 feet in the air, and guitar amps were occasionally stowed away under the stage. These changes required the use of a monitor console—a separate mixer used for the sole purpose of creating multiple monitor mixes for the performers—to accommodate all the additional inputs as well as create separate mixes for each performer. Today, even the smallest music clubs offer at least two or three separate monitor mixes, and it is not uncommon for local bands to carry their own monitor rig capable of handling four or more mixes. Many national touring acts routinely employ upwards of sixteen stereo mixes, Fig. 37-2.

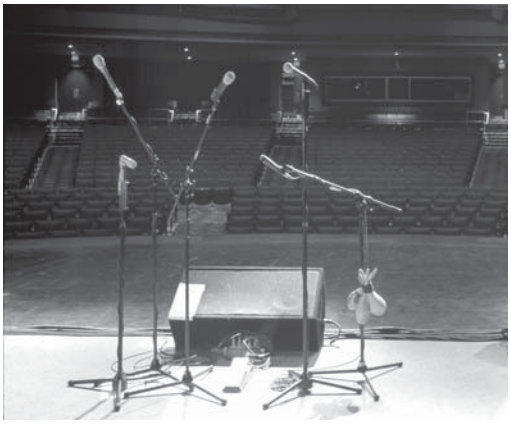


Figure 37-1. Floor loudspeaker monitor wedge. Courtesy Shure Incorporated.

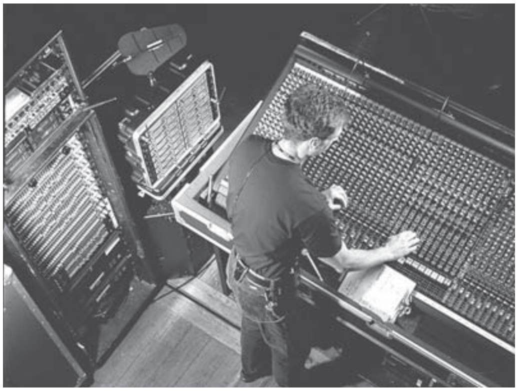


Figure 37-2. Large frame monitor console. Courtesy Shure Incorporated.

The problems created by traditional monitor systems are numerous; the next section examines them in detail. Suffice it to say, a better way to monitor needed to be found. Drummers have used headphones for years to monitor click tracks (metronomes) and loops. Theoretically, if all performers could wear headphones, the need for monitor wedges would be eliminated. Essentially, headphones were the first personal monitors—a closed system that doesn't affect or depend on the monitoring requirements of the other performers. Unfortunately, they tend to be cumbersome and not very attractive. The adoption of transducer designs from hearing aid technology allows performers to use earphones, essentially headphones reduced to a size that fit comfortably in the ear. Professional musicians, including Peter Gabriel and The Grateful Dead, were among the first to employ this new technology. The other major contribution to the development of personal monitors is the growth of wireless microphone systems. Hardwired monitor systems are fine for drummers and keyboardists that stay relatively stationary, but other musicians require greater mobility. Wireless personal monitor systems, essentially wireless microphone systems in reverse, allow the performer complete freedom of movement. A stationary transmitter broadcasts the monitor mix. The performer wears a small receiver to pick up the mix. The first personal monitor systems were prohibitively expensive; only major touring acts could afford them. As with any new technology, as usage becomes more widespread, prices begin to drop. Current personal monitor systems have reached a point where they are within many performers' budgets.

37.2 Personal Monitor System Advantages

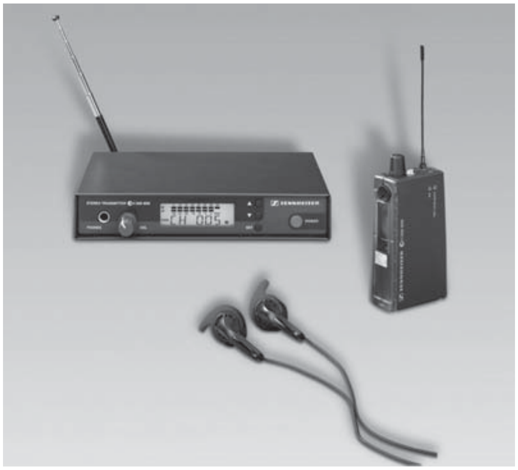
The traditional, floor wedge monitor system is fraught with problems. Performers, especially singers, find it difficult to hear things clearly. Feedback is a constant and annoying issue. And the monitor engineer forever battles to keep up with the needs of the individual performers. Anyone who has performed live has probably dealt with a poor monitor system, but even a great system has many limitations due to the laws of physics, and those laws bend for no one. The concept of in-ear monitoring rose from the desire to create an onstage listening experience that could overcome the limitations imposed by a traditional floor monitor system.

Many parallels exist between personal monitors and a traditional floor wedge setup. The purpose of any monitor system is to allow performers to hear themselves. The sounds to be monitored need to be converted to electronic signals for input to the monitor system. This is usually accomplished via microphones, although in the case of electronic instruments such as keyboards and electronic drums, the signals can be input directly to a mixing console. The various signals are then combined at a mixer, and output to either power amplifiers and loudspeakers or to the inputs of personal monitor systems. Any amount of signal processing, such as equalizers or dynamics processing (compressors, limiters, etc.) can be added inbetween. A hardwired personal monitor system is similar (in signal flow terms) to a traditional wedge system, since the belt pack is basically a power amplifier, and the earphones are tiny loudspeakers. A wireless personal monitor system, however, adds a few more components, specifically a transmitter and receiver, Fig. 37-3. From the output of the mixer, the audio signal goes to a transmitter, which converts it to a radio frequency (RF) signal. A belt-pack receiver, worn by the performer, picks up the RF signal and converts it back to an audio signal. At this stage the audio is then amplified and output to the earphones.

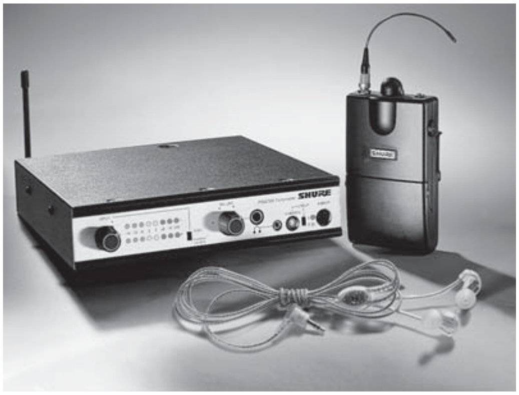
The term personal monitors is derived from several factors, but basically revolves around the concept of taking a monitor mix and tailoring it to each performer's specific needs, without affecting the performance or listening conditions of the others. The concept is broader than that of in-ear monitoring, which states where the monitors are positioned, but gives no further information on the experience.

The four most prominent benefits when using them are:

- Improved sound quality.
- Portability.
- Onstage mobility.
- Personal control.



A. Sennheiser Evolution 300 series IEM.
Courtesy Sennheiser Electronic Corporation.



B. Shure PSM 700 system.
Courtesy Shure Incorporated.

Figure 37-3. Two wireless personal monitor systems.

37.3 Sound Quality

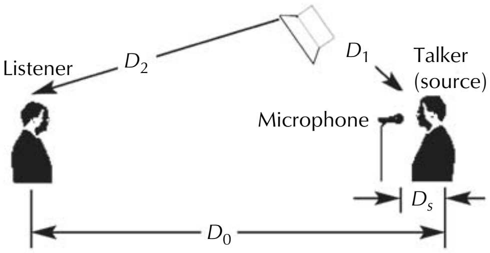
There are several factors that, when taken as a whole, result in improved sound quality with personal monitor systems. These factors include adequate volume for the performers, gain-before-feedback, hearing conservation, reduced vocal strain, and less interference with the audience mix.

37.3.1 Adequate Volume

The most common request given to monitor engineers is "Can you turn me up?" (Sometimes not phrased quite so politely.) Unfortunately, it is not always quite that simple. Many factors can limit how loud a signal can be brought up when using traditional floor monitors: size of the power amplifiers, power handling of the speakers, and most importantly, potential acoustic gain (see Gain-Before-Feedback below). Another factor that makes hearing oneself difficult is the noise level onstage. Many times, vocalists rely solely on stage monitors, unlike guitarists, bassists, and keyboardists whose instruments are generally amplified to begin with. Drummers, of course, are acoustically loud without amplification. Volume wars are not uncommon as musicians struggle to hear themselves over the ever-increasing din. The clarity of the vocals is often obscured as other instruments are added to the monitor mix, which becomes increasingly necessary if fewer mixes are available. Keyboards, acoustic guitars, and other instruments that rely on the monitors often compete with the vocals for sonic space. A personal monitor system, which isolates the user from crushing stage volumes and poor room acoustics, allows the musician to achieve a studiolike quality in the onstage listening experience. Professional, isolating earphones, when used properly, provide more than 20 dB of reduction in background noise level. The monitor mix can then be tailored to individual taste without fighting against otherwise uncontrollable factors.

37.3.2 Gain-Before-Feedback

More amplification and more loudspeakers can be used to achieve higher monitoring levels with traditional stage wedges, but eventually the laws of physics come into play. The concept of gain-before-feedback relates to how loud a microphone can be turned up before feedback occurs. Closely related is PAG, or potential acoustic gain. The PAG equation is a mathematical formula that can be used to predict how much gain is available in a sound system before reaching the feedback threshold, simply by plugging in known factors such as source-to-microphone distance and microphone-to-loudspeaker distance, Fig.37-4. Simply stated, the farther away a sound source is from the microphone, or the closer the microphone is to the loudspeaker, or the farther away the loudspeaker is from the listener, then the less available gain-before-feedback. Now picture a typical stage. The microphone is generally close to the performer's mouth (or instrument); that's good. The microphone is close (relatively) to the monitor loudspeaker; that's bad. The monitor loudspeaker is far (relatively) from the performer's ears; that's also bad. Feedback occurs whenever the sound entering a microphone is reproduced by a loudspeaker and "heard" by the same microphone again. To achieve a decent monitoring level requires quite a bit of available gain. But given the above situation, two major factors drastically reduce the available gain-before-feedback. Compounding the problem is the issue of NOM, or number of open microphones. Every time you double the number of open microphones, the available gain-before-feedback drops by 3 dB. With four open microphones onstage instead of one, the available gain has dropped by 6 dB.



Potential Acoustic Gain

$$PAG = 20\,(\log D_1 - \log D_2 + \log D_0 - \log Ds) - 10\log NOM - 6$$

Figure 37-4. PAG values.

Solutions? The PAG equation assumes omnidirectional microphones, so using cardioid or even supercardioid pattern microphones will help; just don't point them at the speakers. Also, the equation assumes that the sound system has a perfectly flat frequency response. The most commonly employed tool for reducing feedback due to response problems is the graphic equalizer. Since some frequencies will feed back before others, an equalizer allows a skilled user to reduce the monitor system's output of those troublesome frequencies. This technique results in approximately 3–9 dB of additional gain, assuming the microphone position doesn't change. It is common practice for some monitor engineers to attempt to equalize the monitor system to the point where there is no feedback, even with a microphone pointed right into the speaker cone. Unfortunately, the fidelity of the monitor is often completely destroyed in an effort to eliminate feedback using equalizers. Even after equalization has flattened the response of the monitor system, PAG again becomes the limiting factor. At this point, the microphone can't be moved much closer to the sound source, and moving the loudspeaker closer to the performer's ears also makes it closer to the microphone, negating any useful effect on PAG.

Personal monitoring completely removes PAG and gain-before-feedback issues. The "loudspeakers" are now sealed inside the ear canal, isolated from the microphone. With the feedback loop broken, it is possible to achieve as much volume as necessary—which leads to the next topic.

37.3.3 Hearing Conservation

If there's an overriding theme in switching performers to personal monitors, it's so that they can hear themselves better. But it doesn't do any good if eventually they can't hear at all. As mentioned earlier, volume wars on stage are a universal problem. Prolonged exposure to extremely high sound pressure levels can quickly cause hearing to deteriorate. Some performers have taken to wearing ear plugs to protect their hearing, but even the best ear plugs cause some alteration of frequency response. Personal monitors offer a level of hearing protection equal to that of ear plugs, but with the additional benefit of tiny loudspeakers in the plugs. The monitoring level is now in the hands of the performer. If it seems to be too loud, there is no excuse for not turning the monitors down to a comfortable level. The use of an onboard limiter is strongly recommended to prevent high-level transients from causing permanent damage. In larger, complex monitor rigs, outboard compressors and limiters are often employed to offer a greater degree of control and protection.

> *NOTE: Using a personal monitor system does not guarantee that the user will not or cannot suffer hearing damage. These systems are capable of producing levels in excess of 130 dB SPL. Prolonged exposure to these kinds of levels can cause hearing damage. It is up to the individual user to be responsible for protecting his or her own hearing. Please see the section Safe Listening with Personal Monitors for more information.*

Reduced Vocal Strain. Closely related to the volume issue, the ability to hear more clearly reduces vocal strain for singers. In order to compensate for a monitor system that does not provide adequate vocal reinforcement, many singers will force themselves to sing with more power than is normal or healthy. Anyone who makes a living with their voice knows that once you lose it, you lose your livelihood. Every precaution should be taken to protect your instrument, and personal monitors are a key ingredient in helping vocalists continue to sing for years to come. (See Adequate Volume, previously discussed.)

Interference with the Audience Mix. The benefits of personal monitors extend beyond those available to the performer. An unfortunate side-effect of wedge monitors is spill from the stage into the audience area. Although directional at high frequencies, speaker cabinets radiate low-frequency information in a more or less omnidirectional manner. This situation aggravates the already complex task facing the FOH (front-of-house) engineer, who must fight against loud stage volumes when creating the audience mix. The excessive low frequencies coming off the backs of the monitors make the house mix sound muddy and can severely restrict the intelligibility of the vocals, especially in smaller venues. But eliminate the wedges, and the sound clears up considerably.

37.3.4 Portability

Portability is an important consideration for performing groups that travel, and for installations where the sound system or the band performance area is struck after every event. Consider the average monitor system that includes three or four monitor wedges at roughly 40 pounds each, and one or more power amplifiers at 50 pounds—this would be a relatively small monitor rig. A complete personal monitor system, on the other hand, fits in a briefcase. Purely an aesthetic consideration, removing wedges and bulky speaker cables from the stage improves the overall appearance. This is of particular importance to corporate/wedding bands and church groups, where a professional, unobtrusive presentation is as important as sound quality. Personal monitors result in a clean, professional-looking stage environment.

37.3.5 Mobility

Monitor wedges produce a *sweet spot* on stage, a place where everything sounds pretty good, Fig. 37.5. If you move a foot to the left or right, suddenly things do not sound as good anymore. The relatively directional nature of loudspeakers, especially at high frequencies, is responsible for this effect. Using personal monitors, though, is like using headphones—the sound goes where you go. The consistent nature of personal monitors also translates from venue to venue. When using wedges, room acoustics play a large part in the overall quality of the sound. Since professional earphones form a seal against ambient noise, acoustics are removed from the equation. In theory, given the same band with the same members, the monitor settings could remain virtually unchanged, and the mix will sound the same every night.

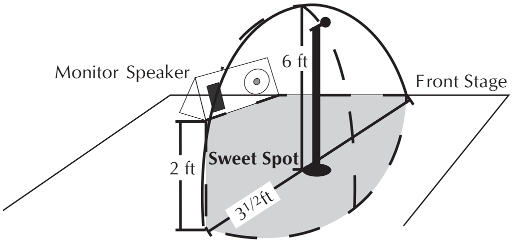


Figure 37-5. Sweet spot created by a wedge monitor loudspeaker.

37.3.6 Personal Control

Perhaps the most practical benefit to personal monitors is the ability for performers to have direct control over what they are hearing. While still relying on the sound engineer to make fine adjustments, personal monitor systems give the performer some ability to make broad adjustments, such as overall volume, pan, or the ability to choose different mixes. If everything in the mix needs to be louder, instead of giving a series of complex hand gestures to the monitor engineer, the performer can raise the overall volume directly from the belt pack.

Many professional systems utilize a *dual-mono* scheme, where the belt pack combines the left and right audio channels of a stereo system and sends the combined signal to both sides of the earphones, Fig. 37-6. The inputs to the system should now be treated as "Mix 1" and "Mix 2" instead of left and right. The balance control on the receiver acts as a mix control, allowing the performer to choose between two mixes, or listen to a combination of both mixes with control over the level of each. Panning to the left gradually increases the level of Mix 1 in both ears, while reducing the level of Mix 2, and vice versa. This feature is referred to by different names, such as MixMode™ (Shure) or FOCUS (Sennheiser), but the function is basically the same. Less expensive, mono-only systems can offer a similar type of control by providing multiple inputs at the transmitter, with a separate volume control for each. Consequently, the transmitter should be located near the performer for quick mix adjustments.

Putting a small, outboard mixer, such as the Shure P4M, near the performer increases the amount of control, Fig. 37-7. By giving control of the monitor mix to the performer, the sound engineer can spend more time concentrating on making the band sound good for the audience instead of worrying about making the band happy.

The cost of transitioning to personal monitors has recently dropped dramatically. A basic system costs as much, if not less than, a typical monitor wedge, power amplifier, and graphic equalizer combination. Expanding a system is also more cost effective. When providing additional wedges for reproducing the same mix, a limited number can be added before the load on the amplifier is too great, and another amp is required. With a wireless personal monitor system, however, the number of receivers monitoring that same mix is unlimited. Additional receivers do not load the transmitter, so feel free to add as many receivers as necessary without adding more transmitters. For bands that haul their own PA, transportation costs may be reduced as well. Less gear means a smaller truck, and possibly one less roadie.

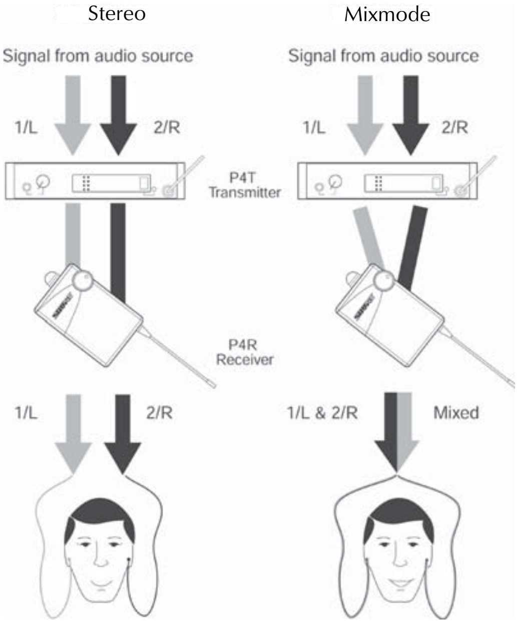


Figure 37-6. How dual-mono works. Courtesy Shure Incorporated.

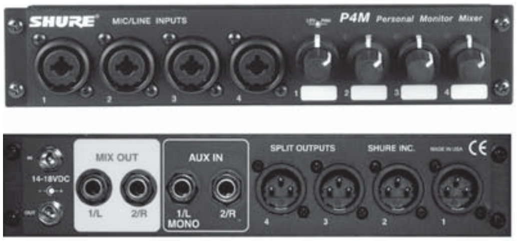


Figure 37-7. Shure P4M outboard mixer. Courtesy Shure Incorporated.

37.4 Choosing a System

Given the personal nature of in-ear monitoring, choosing the right system is an important step. Several choices are available. Present as well as future needs should be taken into account before making an investment.

Personal monitor systems come in two basic varieties—wireless or hardwired. A hardwired system requires the performer be tethered to a cable, which is not necessarily a negative. Drummers and keyboard players who stay relatively stationary, or even backup singers, can take advantage of the lower cost and greater simplicity of a hardwired personal monitor system. Simply connect the monitor sends to the inputs of the hardwired system and dial up a mix. Hardwired systems also work worldwide without the hassle of finding clear frequencies or dealing with local wireless laws and codes. Lastly, if several performers require the same mix, hardwired systems with sufficiently high input impedance can be daisy-chained together without significant signal loss. Alternately, a distribution amplifier could be used to route a signal to multiple hardwired systems. A distribution amplifier takes a single input and splits it to multiple outputs, often with individual level control.

Wireless equipment, by nature, requires special considerations and attention to detail. But the advantages many times outweigh the increased cost and complexity. One of the main benefits of personal monitors is a consistent mix no matter where the performer stands; going wireless allows full exploitation of this advantage. Additionally, when several performers require the same mix, hooking them up is even easier. As many wireless receivers as necessary can monitor the same mix with no adverse effects.

Secondly, consider the travel requirements, if any, of the users. Most wireless equipment, whether it is a microphone system or personal monitors, transmits on unused television channels. Since these unoccupied channels will be different in every city, it is imperative that appropriate frequencies are chosen. For a group that performs only in one metropolitan area, or for a permanent installation, once a good frequency is chosen, there should be no need to change it. However, for touring acts, the ability to change operating frequencies is essential.

The following important specifications for selecting wireless microphones also apply when selecting a personal monitor system:

- Frequency range.
- Tuning range (bandwidth).
- Number of selectable frequencies.
- Maximum number of compatible frequencies.

37.5 Configuring a Personal Monitor System

Choosing the proper system requires some advance planning to determine the monitoring requirements of the situation. At a minimum, the three questions below require answers:

- How many mixes does the situation require.
- Will the monitor mix be stereo or mono.
- How many monitor mixes can be supplied by the mixing console.

This information directly relates to the equipment needed to satisfy the in-ear monitoring requirements of the performers. The following example details the thought process involved in deciding how to configure a system.

37.5.1 How Many Mixes Are Required?

The answer to this question depends on how many performers there are, and their ability to agree on what they want to hear in the monitors. For example, typical rock band instrumentation consists of drums, bass, guitar, keys, lead vocal, and two backup vocals provided by the guitar player and keyboardist. In a perfect world, everyone would want to listen to the same mix, so the answer to this question would be one mix. However, most real-world scenarios require more than one monitor mix. An inexpensive configuration uses two mixes, one consisting of vocals, the other of instruments. Using a system that features dual-mono operation, the performers individually choose how much of each mix they wish to hear, Fig. 37-8. This scenario is a cost-effective way to get into personal monitors, yet still requires a fairly good degree of cooperation among band members.

Another scenario gives the drummer a separate mix, Fig 37-9. This option works well for two reasons:

1. Drummers, in general, will want to hear considerably more drums in the monitors than other band members.
2. For bands who perform on small stages the drums are so loud that they are easily heard acoustically (with no additional sound reinforcement). Therefore, drums may not even be necessary in the other mixes. Now there are three mixes—the vocal mix, the instruments (minus drums), and the drummer's mix.

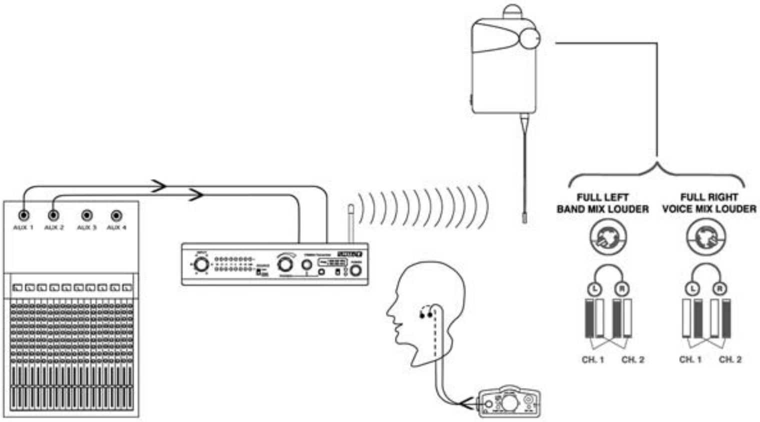

AUX 1
AUX 2
AUX 3
AUX 4
FULL LEFT
BAND MIX LOUDER
FULL RIGHT
VOICE MIX LOUDER
L
R
L
R
CH. 1
CH. 2
CH. 1
CH. 2

Figure 37-8. Two mixes, dual-mono.

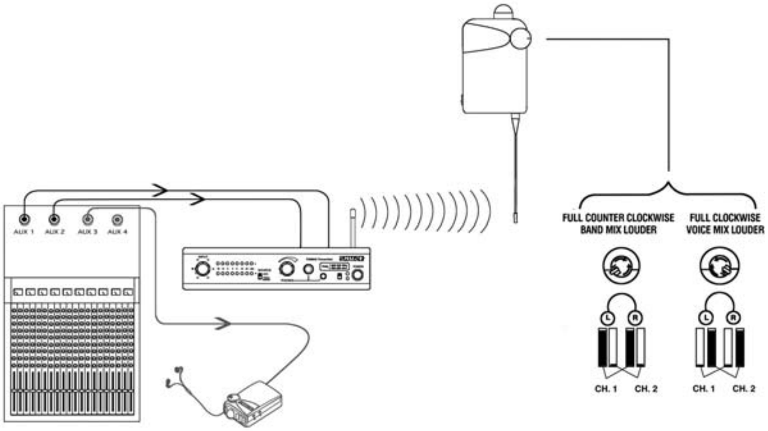

AUX 1
AUX 2
AUX 3
AUX 4
FULL COUNTER CLOCKWISE
BAND MIX LOUDER
FULL CLOCKWISE
VOICE MIX LOUDER
L
R
L
R
CH. 1
CH. 2
CH. 1
CH. 2

Figure 37-9. Three mixes.

Up to this point, it is assumed that the vocalists are able to agree on a mix of the vocal microphones. While forcing singers to share the same mix encourages a good vocal blend, this theory commonly falls apart in practice. Often, separating out the lead vocalist to an individual mix will address this issue, and this can be handled in one of two ways. First, place some of the backup vocal mics in the instruments mix, and adjust the vocal mix to satisfy the lead singer, even if that means adding some instruments to the vocal mix. This scenario results in:

- An individual mix for the lead singer.
- A mix for the guitarist and keyboardist that includes their vocals.
- A drum mix (at this point the bass player can drop in wherever he or she wants, often on the drummer's mix).

The second option is to create a fourth mix for the lead singer, without affecting the other three. This configuration allows the guitarist and keyboardist to retain control between their vocals and instruments, while giving the lead singer a completely customized mix. Does the bass player need a separate mix? That is number five. Adding a horn section? That could easily be a sixth mix. More mixes can be added until one of two limitations is reached; either the mixer runs out of outputs, or the maximum number of compatible frequencies for the wireless monitor system has been reached.

37.5.2 Stereo or Mono?

Most personal monitor systems allow for monitoring in either stereo or mono. At first glance, stereo may seem the obvious choice, since we hear in stereo, and almost every piece of consumer audio equipment offers at least stereo, if not multichannel surround, capabilities. While it may not be applicable to all situations, especially with a limited number of mixes available, a monitor mix created in stereo can more accurately re-create a realistic listening environment. We spend our entire lives listening in stereo; logically, a stereo monitor mix increases the perception of a natural sound-stage. Monitoring in stereo can also allow for lower overall listening levels. Imagine a group with two guitar players sharing the same mix. Both instruments are occupying the same frequency spectrum, and in order for each guitarist to hear, they are constantly requesting their own level be turned up. When monitoring in mono, the brain differentiates sounds based only on amplitude and timbre. Therefore, when two sounds have roughly the same timbre, the only clue the brain has for perception is amplitude, or level. Stereo monitoring adds another dimension, localization. If the guitars are panned, even slightly, from center, each sound occupies its own "space." The brain uses these localization cues as part of its perception of the sound. Research has shown that if the signals are spread across the stereo spectrum, the overall level of each signal can be lower, due to the brain's ability to identify sounds based on their location.

Stereo, by its very nature, requires two channels of audio. What this means for personal monitor users is two sends from the mixer to create a stereo monitor mix —twice as many as it takes to do a mono mix, Fig. 37-10. Stereo monitoring can rapidly devour auxiliary sends; if the mixer has four sends, only two stereo mixes are possible, versus four mono. Some stereo transmitters can be operated in a dual-mono mode, which provides two mono mixes instead of one stereo. This capability can be a great way to save money. For situations that only require one mix, such as solo performer, mono-only systems are another cost-effective option. Strongly consider a system that includes a microphone input that will allow the performer to connect a microphone or instrument directly to the monitor system, Fig, 37-3.

37.5.3 How Many Mixes Are Available from the Mixing Console?

Monitor mixes are typically created using auxiliary sends from the front-of-house (audience) console, or a dedicated monitor console if it's available. A typical small-format console will have at least four auxiliary sends. Whether or not all these are all available for monitors is another matter. Aux sends are also used for effects (reverb, delay, etc.). At any rate, *available auxiliary sends are the final determinant for the number of possible monitor mixes.* If the answer to question 1 (number of required mixes) is greater than the answer to question number 3 (number of mixes available), there are two options: reconfigure the required monitor mixes to facilitate sharing mixes with the existing mixer, or get a new mixer.

37.5.4 How Many Components Are Needed?

After answering the above questions, plug the numbers into the following equations to determine exactly how many of each component are needed and choose a system that can handle these requirements.

Stereo Mixes:
Number of transmitters = number of desired mixes.
Number of aux sends = 2 (number of transmitters),
(ex. 4 mixes requires 4 transmitters and 8 aux sends).

Dual-Mono Mixes:
Number of transmitters = (number of desired mixes)/2
Number of required aux sends = 2(number of transmitters) (ex. 4 mixes requires 2 transmitters and 4 aux sends).

Mono Mixes:
Number of transmitters = number of desired mixes.
Number of aux sends = number of transmitters (ex. 4 mixes requires 4 transmitters and 4 aux sends).

Number of receivers = number of performers

37.6 Earphones

37.6.1 Earphone Options

The key to successful personal monitoring lies in the quality of the earphone. All the premium components in the monitoring signal path will be rendered ineffective by a low-quality earphone. A good earphone must combine full-range audio fidelity with good isolation, comfort, and inconspicuous appearance. The types of earphones available include inexpensive Walkman®-type ear-buds, custom molded earphones, and universal earphones. Each type has its advantages and disadvantages. While relatively affordable, ear-buds have the poorest isolation, are not really designed to withstand the rigors of a working musician's environment, and are likely to fall out of the ear. On the other end of the spectrum, custom molded earphones offer exceptional sound quality and isolation, a considerably higher price tag,

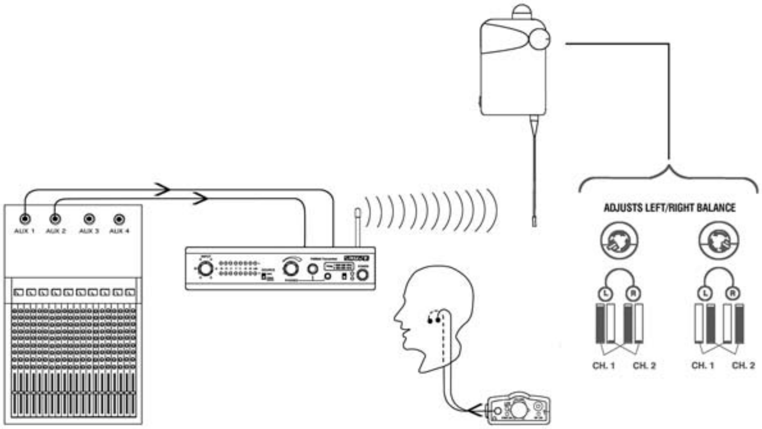


Figure 37-10. One stereo mix.

and are difficult to try before buying since they are made specifically for one person's ears. The procedure for getting custom molds involves a visit to an audiologist. The audiologist makes an impression of the ear canals by placing a dam inside the ear to protect the eardrum, and fills them with a silicone-based material that conforms exactly to the dimensions of the ear canal. The impressions are then used to create the custom molded earphones. Another visit to the audiologist is required for a final fitting. Manufacturers of custom molded earphones include Ultimate Ears, Sensaphonics, and Future Sonics, Fig. 37-11

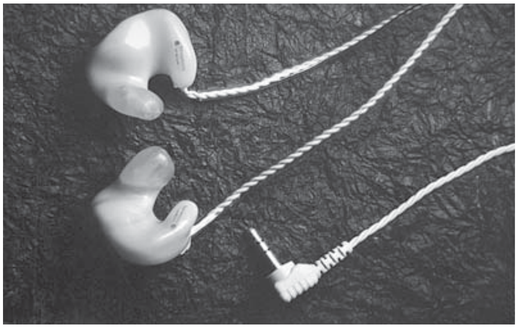


Figure 37-11. Custom molded earphones. Courtesy Sensaphonics.

A third type of earphone is the universal fit, Fig 37-12. Universal earphones combine the superior isolation and fidelity of custom molded designs with the out-of-the-box readiness of ear-buds. The universal nature of this design is attributed to the interchangeable sleeves that are used to adapt a standard size earphone to any size and shape of ear canal. This design allows the user to audition the various sleeves to see which works best, as well as being able to demo the earphones before a purchase is made. The different earphone sleeve options include foam, rubber flex sleeves, rubber flange tips, and custom molded. The foam sleeves resemble regular foam earplugs, but with a small hole in the center of the foam lined with a tube of plastic. They offer excellent isolation and good low-frequency performance. On the downside, they eventually get dirty and worn, and need to be replaced. Proper insertion of the foams also takes longer—relative to the other options—since the earphone needs to be held in place while the foam expands. For quick insertion and removal of the earphones, flexible rubber sleeves may be a good choice. Made of soft, flexible plastic, flex sleeves resemble a mushroom cap and are usually available in different sizes. While the seal is usually not as tight as with the foams, rubber sleeves are washable and reusable. The triple-flange sleeves have three rings (or flanges) around a central rubber tube. They are sometimes referred to as Christmas trees based on their shape. The pros and cons are similar to that of the flex sleeves, but they have a different comfort factor that some users may find more to their liking. The fourth, and most expensive, option is custom sleeves. The custom sleeves combine the relative ease of insertion and permanency of flex sleeves with the superior (depending on the preference of the user) isolation of the foams. The process for obtaining custom sleeves for universal earphones is very similar to that of getting custom molded earphones; a visit to an audiologist is required to get impressions made. Custom sleeves also give the user many of the same benefits as custom molded earphones, but usually at a lower cost, and with the added benefit of being able to interchange earphones with the sleeves if they get lost, stolen, or are in need of repair. A final option, for users of the ear-bud type of earphone, is a rubber boot that fits over the earphone. This option typically has the poorest isolation.

Figure 37-12. Shure SCL5 universal earphone. Courtesy Shure Incorporated.

If there is ever a problem with a universal earphone, another set can be substituted with no negative repercussions. A custom molded earphone does not allow for this kind of versatility; if one needs repair, the only alternative is to have a backup to use in the interim.

> *IMPORTANT NOTE: There are several brands of custom molded earplugs with internal filters that have relatively flat frequency response and different levels of attenuation. Although it may be physically possible to make universal earphones fit into the plugs with the filter removed, this is not advised. The location of the earphone shaft in the ear canal is crucial to obtaining proper frequency response, and most earplugs will prevent them from getting in the proper position. Once again, custom molded earplugs are NOT an acceptable alternative to custom sleeves.*

37.6.2 Earphone Transducers

The internal workings of earphones vary as well. There are two basic types of transducer used in earphone design—dynamic and balanced armature.

The dynamic types work on the same principle as dynamic microphones or most common loudspeakers. A thin diaphragm is attached to a coil of wire suspended in a magnetic field. Diaphragm materials include Mylar (in the case of dynamic microphones) or paper (for loudspeakers). As current is applied to the coil, which is suspended in a permanent magnetic field, it vibrates in sympathy with the variations in voltage. The coil then forces the diaphragm to vibrate, which disturbs the surrounding air molecules, causing the variations in air pressure we interpret as sound. The presence of the magnet-voice coil assembly dictates a physically larger earphone. Dynamic transducers are commonly used in the ear-bud types, but recent technological advances have allowed them to be implemented in universal designs. They are also found in some custom molded earphones.

Originally implemented in the hearing aid industry, the balanced armature transducer combines smaller size with higher sensitivity. A horseshoe-shaped metal arm has a coil wrapped around one end and the other suspended between the north and south poles of a magnet. When alternating current is applied to the coil, the opposite arm (the one suspended in the magnetic field) is drawn towards either pole of the magnet, Fig. 37-13. The vibrations are then transferred to the diaphragm, also known as the *reed*, usually a thin layer of foil. Balanced armature transducers are similar to the elements used in controlled magnetic microphones. In addition to the increased sensitivity, they typically offer better high-frequency response. Achieving a good seal between the earphone and the ear canal is crucial to obtaining proper frequency response.

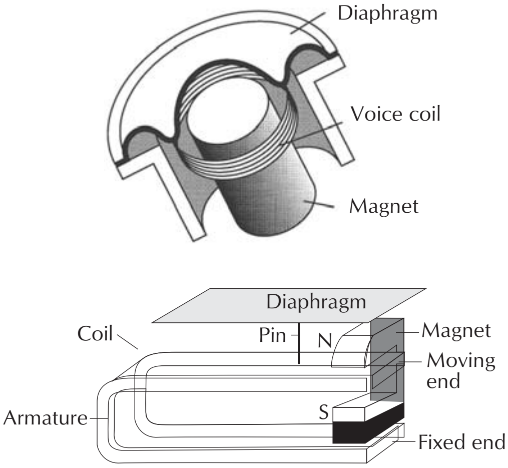


Figure 37-13. Dynamic and balanced armature transducer. Courtesy Shure Incorporated.

A further subdivision occurs with the use of multiple transducers. Dual transducer (*dual driver*) earphones are the most common. Another example of a loudspeaker with dual-transducer design is one with a horn (or tweeter) for high-frequency reproduction and a woofer for low-frequency sounds. The frequency spectrum is divided in two by a crossover network. Each driver only has to reproduce the frequency range for which it has been optimized. Dual driver earphones work on a similar principle—each earphone contains a tweeter and a woofer optimized for high- and low-frequency performance, respectively. Additionally, a passive crossover is built into the cable to divide the audio signal into multiple-frequency bands. The end result is usually much better low end, as well as extended high-frequency response. The additional efficiency at low frequencies may be of particular interest to bassists and drummers. A few companies have introduced triple-driver earphones, and hybrid earphones that combine both dynamic and balanced armature transducers in a single earphone.

37.6.3 The Occluded Ear

One final note for users who are new to earphones. When the ear canal is acoustically sealed (occluded), the auditory experience is different from normal listening. For those performers who have spent many years using traditional floor monitors, an adjustment period may be necessary. A common side effect for vocalists is under-singing. The sudden shock of hearing oneself without straining causes some vocalists to sing softer than they normally would, making it difficult for the FOH engineer to get the vocals loud enough in the house mix. Remember, the FOH engineer is still fighting the laws of PAG, so singers still needsto project.

Another side effect of the occluded ear is a buildup of low frequencies in the ear canal. Sealing off the ear canal such as with an earplug, causes the bones of the inner ear to resonate due to sound pressure levels building up in the back of the mouth. This resonance usually occurs below 500 Hz and results in a hollow sound that may affect vocalists and horn players. Recent studies have shown, however, that ear molds that penetrate deeper into the ear canal (beyond the second bend) actually reduce the occlusion effect. The deeper seal reduces vibration of the bony areas of the ear canal.

37.6.4 Ambient Earphones

Some users of isolating earphones complain of feeling closed off or too isolated from the audience or performance environment. While isolating earphones provide the best solution in terms of hearing protection, many performers would appreciate the ability to recover some natural ambience. There are several ways in which this can be accomplished, the most common being ambient microphones. Ambient microphones are typically placed at fixed locations, nowhere near the listener’s ear, and the levels are controlled by the sound engineer instead of the performer. Additionally, the directional cues provided by ambient microphones (assuming a left/right stereo pair) are dependent on the performer facing the audience. If the performer turns around, the ambient cues will be reversed.

More natural results can be obtained by using a newer technology known as *ambient earphones*. An ambient earphone allows the performer, by either acoustic or electronic means, to add acoustic ambience to the in-ear mix. Passive ambient earphones have a port, essentially a hole in the ear mold, which allows ambient sound to enter the ear canal. While simple to implement, this method offers little in the way of control and could potentially expose the user to dangerous sound pressure levels. Active ambient earphones use minuscule condenser microphones mounted directly to the earphones. The microphones connect to a secondary device that provides the user with a control to blend the desired amount of ambience into the personal monitor mix. Since these microphones are located right at the ear, directional cues remain constant and natural. Ambient earphones not only provide a more realistic listening experience, but also ease between-song communication amongst performers.

37.7 Applications for Personal Monitors

Configuring personal monitor systems and making them work is a relatively simple process, but the ways in which they can be configured are almost limitless. This section takes a look at several typical system set-up scenarios. Personal monitor systems are equally useful for performance and rehearsal, and their benefits extend from small nightclub settings to large arena tours to houses of worship.

37.7.1 Rehearsals

For groups that already own a mixer, implementing a system for rehearsals is a simple process. There are a number of ways to get signal into the system, depending on how many mixes are necessary. To create a simple stereo mix, simply connect the main outputs of the mixer directly to the monitor system inputs. (Note that this works just as well for mono systems). Auxiliary sends can also be used if separate mixes are desired. For bands that carry their own PA system (or at least their own mixer), this method allows them to create a monitor mix during rehearsal, and duplicate it during a performance. No adjustment needs to be made for the acoustic properties of the performance environment.

37.7.1.1 Performance, Club/Corporate/Wedding Bands—No Monitor Mixer

The majority of performing groups do not have the benefit of a dedicated monitor mixer. In this situation, monitor mixes are created using the auxiliary sends of the main mixing console. The number of available mixes is limited primarily by the capabilities of the mixer. At a basic level, most personal monitor systems can provide at least one stereo or two mono mixes. Therefore, any mixer employed should be capable of providing at least two dedicated, prefader auxiliary sends. Prefader sends are unaffected by changes made to the main fader mix. Postfader sends change level based on the positions of the channel faders. They are usually used for effects. Although postfader sends can be used for monitors, it can be somewhat distracting for the performers to hear level changes caused by fader moves.

For users that only have two auxiliary sends available, the best choice is a system that allows a dual-mono operating mode, since this allows for the most flexibility. Hookup is straightforward—just connect Aux Send 1 of the console to the left input and Aux Send 2 to the right input. (Use Aux 3 and 4 if those are the prefader sends—all consoles are different!) Then, depending on who is listening to what, create the mixes by turning up the auxiliary sends on the desired channels. A few common two-mix setups are listed below.

Each performer can choose which mix they want to listen to by adjusting the balance control on the receiver. Be sure the receivers are set for dual-mono operation, or each mix will only be heard on the left or right side, but not in both ears. Also remember that any number of receivers can monitor the same transmitter.

Some performers may prefer to listen to the house mix, so they can monitor exactly what the audience is hearing. Keep in mind that this may not always produce the desired results. Rarely will what sounds good in the ear canal sound equally as good through a PA system in a less-than-perfect acoustic environment. Many times, a vocal that seems to blend just right for an in-ear mix will get completely lost through the PA, especially in a small room when live instruments are used. This technique may be appropriate for electronic bands, where the majority of instruments is input directly to the mixer. The only sound in the room is that created by the sound system.

The more auxiliary sends a console has, the greater number of monitor mixes that is possible. See Tables 37-1 to 37-3 for more examples.

Table 37-1. Three Monitor Mixes (MixMode™)

Option 1

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
Vocal mix	Band mix	Dedicated drum mix	Unused

Option 2

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
Lead vocal	Everything else	Dedicated drum mix	Unused

Option 3

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
Front mix	Backline mix	"Ego" mix (bandleader gets whatever he or she wants)	Unused

Table 37-2. Four Monitor Mixes (MixMode™—Using Only Aux Sends and PSM Loop Jacks)

Option 1

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
Vocal mix	Band mix	Horn mix	Band mix (looped from PSM Right Loop Out Jack)

Table 37-3. Four Monitor Mixes (MixMode™)

Option 1

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
Lead vocalist's mix	Guitarist's mix	Bassist's mix	Drummer's mix

Option 2

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
Vocal mix	Band mix	Horn mix	Vocal/band mix

Option 3

Aux 1 Out (PSM 1 Left)	Aux 2 Out (PSM 1 Right)	Aux 3 Out (PSM 2 Left)	PSM 2 Right
"EGO" mix (lead vocal/ instrument only)	"Ego" mix (everything else)	Band mix	Dedicated drum mix

37.7.1.2 Club Level Bands—With Monitor Console

At this point, a typical small format FOH console has reached its limit for monitoring purposes. Bands that have graduated to the next level of performance (larger, more prestigious clubs and theaters or small tours) may find themselves in a position to take advantage of a dedicated monitor console. Most monitor boards are capable of providing at least eight mono (or four stereo) mixes. It now becomes practical for each band member to have his or her own dedicated mix. System hookup is again very simple—the various mix outputs from the monitor console are connected directly to the personal monitor system. Stereo monitoring is a much more viable option due to the large number of mixes available, as well as the presence of a skilled monitor engineer to adjust the mixes to the point of perfection.

Some performers even carry their own monitor console. Due to the consistent nature of personal monitors, a band with the same instrumentation and performers for every show can leave the monitor mix dialed-in on its console. Since venue acoustics can be completely disregarded, a few minor adjustments are all that is typically necessary during sound check.

A personal monitor mixer can also be used to augment the monitor console, if the performer desires some personal control over what is heard. In the past, drummers or keyboard players would use a small mixer and Y-cables to submix their instruments for in-ear monitors. A mixer with built-in mic splitting capability, such as the Shure P4M Personal Monitor Mixer, Fig. 37-7, can be used in the same capacity, but without the need for Y-cables.

37.7.1.3 Professional Touring System

When budget is no longer a consideration, personal monitoring can be exploited to its fullest capabilities. Many systems used by professional artists on large-scale tours often employ greater than sixteen stereo mixes.

A completely separate, totally personalized mix is provided for every performer onstage. Large frame monitor consoles are a requirement. For example, to provide sixteen stereo mixes requires a monitor console with thirty two outputs.

Effects processing is generally employed to a much larger extent than with a smaller system.

When operating a large number of wireless personal monitor systems, R-related issues become much more important. Frequency coordination must be done carefully to avoid interaction between systems as well as outside interference. Depending on the extent of the touring, a frequency agile system is desirable, if not required. Proper antenna combining, to reduce the number of transmitter antennas in close proximity, is a necessity. Directional antennas may also be used to increase range and reduce the chances of drop-outs due to multipath interference.

37.7.2 Mixing for Personal Monitors

Mixing for personal monitors may require a different approach than one used for a traditional monitor system. Often, the requirements for the performers are very involved, and can require a greater degree of attentiveness from the monitor engineer. In particular, many small nightclub sound systems typically provide monitors for the sole purpose of vocal reinforcement. An in-ear monitor system, due to its isolating nature, usually demands other sound sources be added to the monitor mix, especially if the instrumentalists choose to reduce their overall stage volume. Some performers may prefer a more active mix, such as hearing solos boosted, or certain vocal parts emphasized. This luxury usually requires a dedicated monitor console and sound engineer. The FOH engineer has enough responsibility mixing for the audience, and generally only alters the monitor mix on request from the performers. In most situations, except for the upper echelon of touring professionals, this approach is perfectly acceptable and still far superior to using wedges.

For performers who are mixing for themselves, there are other considerations. One of the advantages of having a professional sound engineer or monitor engineer is years of experience in mixing sound. This skill cannot be learned overnight. For bands that are new to personal monitors, there is a strong temptation to try to create a CD-quality mix for in-ears. While this is certainly possible with a trained sound engineer and the right equipment, it is unlikely that someone unfamiliar with the basic concepts behind mixing will be able to successfully imitate a professional mix.

A common mistake made by in-ear monitor novices is to put everything possible into the mix. Here's an alternative to the everything-in-the-mix method:

1. Put the earphones on and turn the system on. DO NOT put any instruments in the mix yet.
2. Try to play a song. While performing, determine what instruments need more reinforcement.
3. Begin bringing instruments into the mix, one at a time. Usually, vocals come first since those are often the only unamplified instruments onstage.

4. Only turn things up as loud as necessary, and resist the temptation to add instruments to the mix that can be heard acoustically.

A note on monitor mixing: performers now have an unprecedented level of personal control over what they are hearing. The temptation to make oneself the loudest thing in the mix is great, but this may not be the best for the situation. Proper blending with the other members of the ensemble will be next to impossible if the mix is skewed too far from reality. Consider big bands that normally play acoustically, or a vocal choir. These types of ensembles create their blend by listening to each other, not just themselves. If the lead trumpet player uses a personal monitor system, and cranks the trumpet up three times louder than everything else, there is no accurate reflection for the musician on whether he or she is playing too loud or too soft. Remember, great bands mix themselves—they don't rely entirely on the sound tech to get it right.

37.7.3 Stereo Wireless Transmission

Many microphones and most circuitry used in the reproduction of audio signals have a bandwidth of 20 kHz (20 Hz–20 kHz) or more. Digital devices operating at a sampling rate of 44.1 kHz have frequency responses extending to 22 kHz; 48 kHz sample rates are flat out to 24 kHz. Many boutique analog devices boast flat response beyond 30 kHz and sometimes 40 kHz. Undoubtedly, these are great advances in the sound reproduction field, unless audio with that kind of bandwidth is sent through the air in the form of a stereo encoded wireless transmission.

Every so often, a performer will complain of an ear mix that just doesn't sound quite right, no matter what adjustments are made to levels or EQ. Sometimes it's a simple image shift; other times it is distortion with no apparent cause. The output of the mixing console sounds fine, as does the headphone output of the transmitter. Changing frequencies, cables, earpieces, and bodypacks makes no difference. Ultimately, respecting the frequency response limitations of stereo wireless transmission is the key to successfully creating stable, good-sounding ear mixes. There are several ways this can be accomplished, but as a general rule, avoid any frequency boosts above 15 kHz. Stereo multiplexed wireless transmission has a limited frequency response of 50 Hz–15 kHz. This frequency response limitation has been in place since the FCC approved stereo multiplexed transmissions (MPX) back in 1961. Audio engineers mixing stereo wireless transmissions for on-stage talent wearing in-ear monitors should be aware of the operating principles of MPX stereo to achieve the desired results at the receiver.

In many cases, switching to mono transmission clears up any wireless anomaly (except for interference) in these types of monitoring systems. However, many users want to monitor in stereo, so being aware of the limitations of MPX encoding will allow for greater talent satisfaction.

Stereo wireless transmitters use a steep cut filter, or brick-wall filter, prior to modulation, centered at 19 kHz to create a safe haven for the required pilot tone. MPX encoders in stereo wireless transmitters use a 19 kHz pilot tone to inform receivers that the transmission is encoded in stereo. If the receiver does not sense a 19 kHz pilot tone, it will only demodulate a mono signal. Moreover, if the 19 kHz pilot tone is not stable, stereo imaging degrades at the receiver. Most importantly, if in-ear monitor receivers do not sense stable 19 kHz pilot tones, they will mute (this is called *tone-key squelch*, a circuit designed to keep the receiver muted when the corresponding transmitter is turned off). Problems are created due to the extensive EQ capabilities of modern mixing consoles, which offer high-frequency shelving equalization from as low as 10 kHz to as high 12, 15, and even 16 kHz. Digital mixing consoles offer parametric filtering that can center on practically any frequency and boost by as much as 18 dB. With a multichannel mixing board, it is easy enough to create a counteractive frequency response at the frequency of interest—19 kHz. In stereo wireless, there are two pieces of information actually being transmitted, the mono or sum signal (left + right) and the difference (left – right) channel, each occupying a 15 kHz-wide swath of spectrum. The 19 kHz pilot tone is centered in between these two signals, Fig. 37-14.

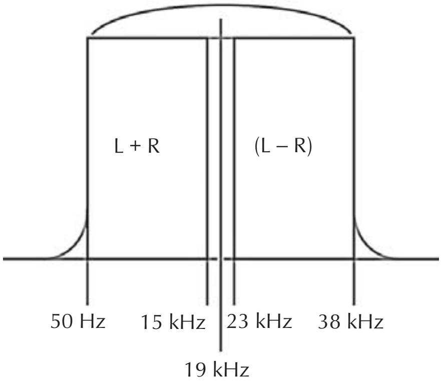


Figure 37-14. Stereo MPX encoding.

The stereo image is restored in the receiver by adding the sum and difference signals to create the left channel, and subtracting them to derive the right channel.

$$(L + R) + (L - R) = 2L$$

$$(L + R) - (L - R) = 2R$$

This system ensures mono compatibility, since the received signal will simply collapse to mono when the pilot tone is lost. Only the $L + R$ sum signal remains.

However, since the 19 kHz pilot tone resides in the audio band, it can easily be compromised by the program material. The result of these high-frequency components getting into the modulator can cause, at best, degradation of stereo separation and distortion, and in worst-case situations, muting of the receiver. Add the high-frequency shelf used in the pre-emphasis curves prior to the companding circuits in stereo transmitters (a form of noise reduction), and it is easy to see how a small high-frequency boost on a channel strip can have a huge effect on what is heard after the RF link. If the audio signal modulates the pilot tone, stereo reception and the resultant sound quality will be poor. If upper harmonics of musical instruments aggravate the $(L - R)$ sidebands (especially in a transient manner—tambourines, triangles, high hats, click tracks, etc.), stereo separation can degrade, frequency response can be compromised, and even dynamic interactions between one channel and another can be detected.

Several simple practices go a long way toward improving stereo transmission:

- Refrain from extreme stereo panning. Instead of panning hard left and right, try the 10 o'clock and 2 o'clock positions.
- Use equalization sparingly prior to stereo transmission for smoother MPX encoding.
- Use $\frac{1}{10}$ of an octave notch filters at 16 kHz on console output busses to increase the slope of the MPX filter. This is the best way to avoid disturbing the pilot tone.

37.7.4 Personal Monitors for Houses of Worship and Sound Contractors

The advantages of using personal monitors extend beyond those of just the performers. The above examples illustrate the benefits to the performer, and from a strictly music industry-oriented point of view. This section will discuss how personal monitors can be a useful tool for the sound contractor, specifically as they apply to modern houses of worship.

Musical performances are rapidly becoming a more prominent part of the worship service. Praise teams and contemporary music groups, while bringing new levels of excitement to traditional church services, also bring with them the problems of an average rock band. Most prominent among these problems are volume wars. Drums naturally tend to be the loudest thing on stage. The guitarist, in order to hear himself better, turns his amplifier up louder. The singers then need more monitor level to compete with the rest of the band. And then the cycle begins again. In any live sound situation, church or otherwise, loud stage volumes can distract from the overall sound in the audience. Try an easy experiment at the next sound check. When the band is satisfied with the monitor mix, turn off the audience PA and just listen to the sound coming off the stage. It's probably loud enough that the main sound system doesn't need to be turned on! To compound matters, the "backwash" off the floor monitors consists primarily of low-frequency information that muddies-up the audience mix. While this situation creates headaches for most sound engineers, it is even worse in the church environment. The majority of Sunday morning service attendees are not looking for an extremely loud rock and roll concert, but in some cases the congregation mix gets this loud just so it can be heard over the stage monitors. If the main system is off, and it's still too loud, what can be done? Turn down the floor monitors and the band complains—not to mention how terrible it will sound.

With the band using personal monitors, these problems evaporate. Traditional floor monitors can be completely eliminated. For part two of our experiment, turn off the stage monitors while the band is playing. Notice how much clearer the audience mix becomes? This is how it would sound if the band were using personal monitors. Also, personal monitors are not just for vocalists. Drummers with in-ear monitors tend to play quieter. When the loudest instrument on stage gets quieter, everything else can follow suit. Some churches take this a step further by using electronic drums, which create little, if any, acoustic noise. Bass, keyboard, and electric guitar can also be taken directly into the mixer if the players are using personal monitors, eliminating the need for onstage amplifiers. The end result is a cleaner, more controlled congregation mix, and musicians can have very loud monitors without affecting the congregation.

Secondly, consider the feedback issue. Feedback occurs when the sound created at the microphone comes out of a loudspeaker, and reenters the microphone. The closer the loudspeaker is to the microphone, the greater the chance for feedback. Eliminating the floor monitor also eliminates the worst possible feedback loop. With the "loudspeakers" sealed inside the ear canal, there is no chance for the signal to reenter the microphone. No equalizer or feedback reducer will ever be as effective as personal monitors at eliminating feedback on the stage.

Many other uses are possible for personal monitors. Choir directors could use them for cues, or to hear the pastor more clearly. Pastors who desire monitor reinforcement of their speech microphones, a sure-fire recipe for feedback, will find this a much better solution. Organists located at the rear of the sanctuary could use them to better hear the choir located up front, or also to receive cues. The advantages extend well beyond the benefits to the performer, and increase the overall quality of the service and the worship experience.

37.8 Expanding the Personal Monitor System

37.8.1 Personal Monitor Mixers

Personal monitoring gives the performer an unprecedented level of control. But for the performer who desires more than simple volume and pan operation, an additional mixer may be implemented. Personal monitor mixers are especially useful for bands who have a limited number of available monitor mixes, or who do not have a monitor engineer, or anyone at all to run sound. In a perfect world, all performers would be happy listening to the exact same mix; in reality, everyone may want to hear something different. A small mixer located near the performers allows them to customize their mix to hear exactly what they desire. Theoretically, any mixer can double as a personal monitor mixer, but most lack one key feature; the input signals need to find their way to the main (FOH) mixer somehow. Large sound systems with separate monitor consoles use transformer-isolated splitters to send the signals to two places, but these are prohibitively expensive for most working bands and small clubs. Y-cables can be used to split microphone signals, but they can get messy and are somewhat unreliable. A few manufacturers produce mixers with integrated microphone splitters. These range from basic four channel mixers with only volume and pan controls for creating a single mix to larger monitor consoles that can provide four or more stereo mixes along with fader control and parametric equalization.

37.8.2 Distributed Mixing

Distributed mixing is the direct result of advances in the area of digital audio networking. By converting analog audio signals to digital, audio can be routed to many locations without degradation or appreciable signal loss. Unlike with analog personal mixers, cabling is far simpler. Typically, analog outputs from a mixing console connect to an analog-to-digital converter. Multiple channels of digital audio can then be routed from the A/D converter to personal mixing stations located by each performer, using a single common Ethernet (Cat-5) cable, thus eliminating a rat's nest of microphone cables or the large, unwieldy cable snakes required for analog audio distribution. Cat-5 cable is inexpensive and readily available. The mixing station provides an analog headphone output that can drive a set of isolating earphones directly, or better yet, connect to either a hardwired or wireless personal monitor system. If nothing else, the personal monitor system offers the advantage of a limiter for some degree of hearing protection, as well as a volume control at the performer's hip. The mixers supplied with most distributed systems do not always have a limiter. Most systems provide eight or sixteen channels of audio, allowing each performer to create his or her own custom mix, independent of other performers and without the intervention of a sound engineer. Note that giving this level of control to the performers will probably require some training in the basics of mixing to be successful (see Creating a Basic Monitor Mix above).

37.8.3 Supplementary Equipment

In-ear monitoring is a different auditory experience from traditional stage monitoring. Since your ears are isolated from any ambient sound, the perception of the performance environment changes. There are several other types of audio products that can be added to a personal monitor system to enhance the experience, or try to simulate a more "live" feel.

37.8.3.1 Drum Throne Shakers

Something performers may miss when making the transition to personal monitors are the physical vibrations created by amplified low-frequency sounds. Drummers and bass players are particularly sensitive to this effect. Although using a dual driver earphone usually results in more perceived bass, an earphone cannot replicate the physical sensation of air moving (sound) anywhere but in the ear canal. Drum shakers exist not to provide audible sound reinforcement, but to re-create the vibrations normally produced by subwoofers or other low-frequency transducers. Commonly found in car audio and cinema applications, these devices mechanically vibrate in sympathy with the musical program material, simulating the air disturbances caused by a loud subwoofer, Fig. 37-15. They can be attached to drum thrones or mounted under stage risers.

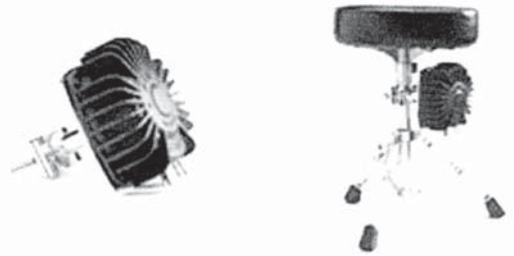


Figure 37-15. Aura Bass Shaker. Courtesy AuraSound, Inc.

37.8.3.2 Ambient Microphones

Ambient microphones are occasionally employed to restore some of the "live" feel that may be lost when using personal monitors. They can be used is several ways. For performers wishing to replicate the sound of the band on stage, a couple of strategically placed condenser microphones can be fed into the monitor mix. Ambient microphones on stage can also be used for performers to communicate with one another, without being heard by the audience. An extreme example (for those whose budget is not a concern) is providing each performer with a wireless lavalier microphone, and feeding the combined signals from these microphones into all the monitor mixes, but not the main PA. Shotgun microphones aimed away from the stage also provide good audience pick-up, but once again, a good condenser could suffice if shotguns are not available.

37.8.3.3 Effects Processing

Reverberant environments can be artificially created with effects processors. Even an inexpensive reverb can add depth to the mix, which can increase the comfort level for the performer. Many singers feel they sound better with effects on their voices, and in-ear monitors allow you to add effects without disturbing the house mix or other performers.

Outboard compressors and limiters can also be used to process the audio. Although many personal monitor systems have a built-in limiter, external limiters will provide additional protection from loud transients. Compression can be used to control the levels of signals with wide dynamic range, such as vocals and acoustic guitar, to keep them from disappearing in the mix. More advanced monitor engineers can take advantage of multiband compression and limiting, which allows dynamics processing to act only on specific frequency bands, rather than the entire audio signal.

In-ear monitor processors combine several of these functions into one piece of hardware. A typical in-ear processor features multiband compression and limiting, parametric equalization, and reverb. Secondary features, such as stereo spatialization algorithms that allow for manipulation of the stereo image, vary from unit to unit.

37.8.4 Latency and Personal Monitoring

An increasing number of devices used to enhance the personal monitor system are digital instead of analog. While the advantages of digital are numerous, including more flexibility and lower noise, any digital audio device adds a measurable degree of *latency* to the signal path, which should be of interest to personal monitor users. Latency, in digital equipment, is the amount of time is takes for a signal to arrive at the output after entering the input of a digital device. In analog equipment, where audio signals travel at the speed of light, latency is not a factor. In digital equipment, however, the incoming analog audio signal needs to be converted to a digital signal. The signal is then processed, and converted back to analog. For a single device, the entire process is typically not more than a few milliseconds.

Any number of devices in the signal path might be digital, including mixers and signal processors. Additionally, the signal routing system itself may be digital. Personal mixing systems that distribute audio signals to personal mixing stations for each performer using Cat-5 cable (the same cable used for Ethernet computer networking) actually carry digital audio. The audio is digitized by a central unit and converted back to analog at the personal mixer. Digital audio snakes that work in a similar manner are also gaining popularity.

Since the latency caused by digital audio devices is so short, the signal will not be perceived as audible delay (or echo). Generally, latency needs to be more than 35 ms to cause a noticeable echo. The brain will integrate two signals that arrive less than 35 ms apart. This is known as the Haas Effect, named after Helmut Haas who first described the effect. Latency is cumulative, however, and several digital devices in the same signal path could produce enough total latency to cause the user to perceive echo.

As discussed, isolating earphones are the preferred type for personal monitors, because they provide maximum isolation from loud stage volume. Isolating earphones, however, result in an effect known as the occluded ear. Sound travels by at least two paths to the listener's ear. The first is a direct path to the ear canal via bone conduction. An isolating earphone reinforces this path, creating a build-up of low frequency information that sounds similar to talking while wearing earplugs. Secondly, the "miked" signal travels through the mixer, personal monitor transmitter and receiver, and whatever other processing may be in the signal path. If this path is entirely analog, the signal travels at the speed of light, arriving at virtually the same time as the direct (bone-conducted" sound. Even a small amount of latency introduced by digital devices, though, causes comb filtering.

Before continuing, an explanation of comb filtering is in order. Sound waves can travel via multiple paths to a common receiver (in this case the ear is the receiver). Some of the waves will take a longer path than others to reach the same point. When they are combined at the receiver, these waves may be out of phase. The resultant frequency response of the combined waves, when placed on a graph, resembles a comb, hence the term *comb filtering*, Fig. 37-16.

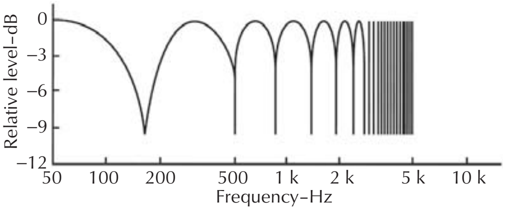


Figure 37-16. Comb filtering.

Hollow is a word often used to describe the sound of comb filtering.

It is generally believed that the shorter the latency, the better. Ultimately, changing the amount of latency shifts the frequency where comb filtering occurs. Even latency as short as 1 ms produces comb filtering at some frequencies. What changes is the frequency where the comb filtering occurs. Lower latency creates comb filtering at higher frequencies. For most live applications, up to 2 ms of delay is acceptable. When using personal monitors, though, total latency should be no more than 0.5 ms to achieve sound quality equivalent to an analog, or zero latency, signal path. While in reality it may be difficult to achieve latency this short, be aware that any digital device will cause some latency. The individual user will have to determine what amount of latency is tolerable. As an alternative, some users report that inverting the polarity of certain input channels, or even the entire mix, improves the sound quality. Keep in mind that comb filtering still occurs, but at frequencies that may be less offensive to the listener.

The degree of latency is generally not more than a few milliseconds, which, as mentioned, will not cause the processed signal to be perceived as an audible delay. The concern for users of in-ear monitors, though, lies primarily with horn players, and occasionally vocalists. When a horn player sounds a note, the vibrations are carried directly to the ear canal via bone conduction. If the microphone signal is subject to digital processing, too much latency can cause comb filtering. The user generally perceives this as a hollow, unnatural sound. Care should be taken to avoid introducing unnecessary processing if comb filtering occurs. Adjusting the delay time in the processor (assuming digital delay is one of the available effects) could also compensate for latency. Alternately, route the effects through an auxiliary bus, rather than right before the monitor system inputs, which will minimize the latency effect by keeping the dry signal routed directly to the monitor system.

37.8.5 Safe Listening with Personal Monitors

No discussion of monitoring systems would be complete without some discussion of human hearing. The brain's ability to interpret the vibrations of air molecules as sound is not entirely understood, but we do know quite a bit about how the ear converts sound waves into neural impulses that are understood by the brain.

The ear is divided into three sections; the outer, middle, and inner ear, Fig. 37-17. The outer ear serves two functions—to collect sound and act as initial frequency response shaping. The outer ear also contains the only visible portion of the hearing system, the pinna. The pinna is crucial to localizing sound. The ear canal is the other component of the outer ear, and provides additional frequency response alteration. The resonance of the ear canal occurs at approximately 3 kHz, which, coincidentally, is right where most consonant sounds exist. This resonance increases our ability to recognize speech and communicate more effectively. The middle ear consists of the eardrum and the middle ear bones (ossicles). This section acts as an impedance-matching amplifier for our hearing system, coupling the relatively low impedance of air to the high impedance of the inner ear fluids. The eardrum works in a similar manner to the diaphragm of a microphone, it moves in sympathy to incoming sound waves, and transfers those vibrations to the ossicles. The last of these bones, the stapes, strikes an oval-shaped window that leads to the cochlea, the start of the inner ear. The cochlea contains 15,000 to 25,000 tiny hairs, known as cilia, which bend as vibrations disturb the fluids of the inner ear. This bending of the cilia sends neural impulses to the brain via the auditory nerve, which the brain interprets as sound.

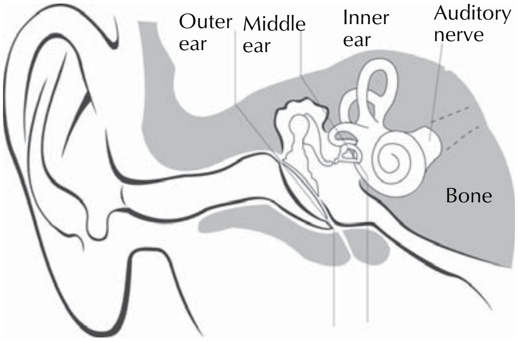


Figure 37-17. Illustration of ear anatomy. Courtesy Shure Incorporated.

Hearing loss occurs as the cilia die. Cilia begin to die from the moment we are born, and they do not regenerate. The cilia that are most sensitive to high frequencies are also the most susceptible to premature damage. Three significant threats to cilia are infection, drugs, and noise. Hearing damage can occur at levels as low as 90 db SPL. According to OSHA (Occupational Safety and Health Administration), exposure to levels of 90 dB SPL for a period of 8 hours could result in some damage. Of course, higher levels reduce the amount of time before damage occurs.

Hearing conservation is important to everyone in the audio industry. As mentioned before, an in-ear monitor system can assist in helping to prevent hearing damage —but it is not foolproof protection. The responsibility for safe hearing is now in the hands of the performer. At this time, there is no direct correlation between where the volume control is set and the sound pressure level present at the eardrum. Here are a few suggestions, though, that may help users of personal monitors protect their hearing.

37.8.5.1 Use an Isolating Earphone

Without question, the best method of protection from high sound pressure levels is to use a high-quality earplug. The same reasoning applies to an in-ear monitor. When using personal monitors, listening at lower levels requires excellent isolation from ambient sound, similar to what is provided by an earplug. Hearing perception is based largely on signal to noise. To be useful, desired sounds must be at least 6 dB louder than any background noise. Average band practice levels typically run 110 dB SPL, where hearing damage can occur in as little as 30 minutes. Using a personal monitor system with a nonisolating earphone would require a sound level of 116 dB SPL to provide any useful reinforcement, which reduces the exposure time to 15 minutes. Inexpensive ear buds, like those typically included with portable MP3 players, offer little, if any, isolation. Avoid these types of earphones for personal monitor applications.

Not all types of isolating earphones truly isolate, either. Earphones based on dynamic drivers typically require a ported enclosure to provide adequate low frequency response. This *port,* a small hole or multiple holes in the enclosure, drastically reduces the effectiveness of the isolation. Note that not all dynamic earphones require ports. Some designs use a sealed, resonating chamber to accomplish the proper frequency response, thus negating the need for ports but preserving the true isolating qualities of the earphone. Earphones that employ a balanced armature transducer, similar to those used in hearing aids, are physically smaller and do not require ports or resonating chambers. In fact, balanced armature-type earphones rely on a good seal with the ear canal to obtain proper frequency response. They can be made somewhat smaller, but are typically more expensive, than their dynamic counterparts.

37.8.5.2 Use Both Earphones!

A distressing, yet increasingly common, trend is only using one earphone and leaving the other ear open. Performers have several excuses for leaving one ear open, the most common is a dislike for feeling removed from the audience, but the dangers far outweigh this minor complaint. First, consider the above example of a 110 dB SPL band practice. One ear is subjected to the full 110 dB, while the other ear needs 116 dB to be audible. Using only one earphone is equivalent to using a nonisolating earphone, except one ear will suffer damage twice as fast as the other. Second, a phenomenon known as *binaural summation*, that results from using both earphones, tricks the ear-brain mechanism into perceiving a higher SPL than each ear is actually subjected to. For example, 100 dB SPL at the left ear and 100 dB SPL in the right ear results in the perception of 106 dB SPL. Using only one earphone would require 106 dB SPL at that ear. The practical difference is potential hearing damage in one hour instead of two. Using both earphones will usually result in overall lower listening levels.

Table 37-4 shows OSHA recommendations for exposure time versus sound pressure level.

Ambient microphones are commonly employed to help overcome the closed-off feeling. An ambient microphone can be a lavalier clipped to the performer and routed directly to the in-ear mix, or a stereo microphone pointed at the audience. The common thread is that they allow the user to control the level of the ambience.

Table 37-4. OSHA Recommended Exposure Time Versus Sound Pressure Level

Sound Pressure Level	Exposure time
90 dB SPL	8 hours
95 dB SPL	4 hours
100 dB SPL	2 hours
105 dB SPL	1 hour
110 dB SPL	30 minutes
115 dB SPL	15 minutes

37.8.5.3 Keep the Limiter On

Unexpected sounds, such as those caused by someone unplugging a phantom-powered microphone or a blast of RF noise, can cause a personal monitor system to produce instantaneous peaks in excess of 130 dB SPL, the equivalent of a gun shot at the eardrum. A brickwalltype limiter can effectively prevent these bursts from reaching damaging levels. *Only use a personal monitor system that has a limiter at the receiver, and do not defeat it for any reason*. A well-designed limiter should not adversely affect the audio quality, as it only works on these unexpected peaks. If the limiter seems to be activating too often, then the receiver volume is probably set too high (read as: unsafe!). Outboard compressors and limiters placed before the inputs of the monitor system are certainly appropriate, but are not a substitute for an onboard limiter, as they cannot protect against RF noise and other artifacts that may occur post-transmitter.

37.8.5.4 Pay Attention to What Your Ears Are Telling You

Temporary threshold shift (TTS) is characterized by a stuffiness, or compressed feeling, like someone stuck cotton in the ears. Ringing (or tinnitus) is another symptom of TTS. Please note, though, that hearing damage may have occurred even if ringing never occurs. In fact, the majority of people who have hearing damage never reported any ringing. After experiencing TTS, hearing may recover. Permanent damage has possibly occurred, though. The effects of TTS are cumulative, so a performer who regularly experiences the above effects is monitoring too loud and hearing damage will occur with repeated exposure to those levels.

37.8.5.5 Have Your Hearing Checked Regularly

The only certain way to know if an individual's listening habits are safe is to get regular hearing exams. The first hearing test establishes a baseline that all future hearing exams are compared against to determine if any loss has occurred. Most audiologists recommend that musicians have their hearing checked at least once a year. If hearing loss is caught early, corrections can be made to prevent further damage.

A frequently asked question about in-ear monitors is: "How do I know how loud it is?" At this time, the only way to develop a useful correlation between the volume knob setting and actual SPL at the eardrum is by measuring sound levels at the eardrum with specially made miniature microphones. A qualified audiologist (not all have the right equipment) can perform the measurements and offer recommendations for appropriate level settings.

Personal monitors can go a long way toward saving your hearing, but only when used properly. Monitoring at lower levels is the key to effective hearing conservation, and this can only be accomplished through adequate isolation. Used correctly, professional isolating earphones, combined with the consultation of an audiologist, offer the best possible solution for musicians interested in protecting their most valuable asset. It cannot be stated strongly enough: a personal monitor system, in and of itself, does not guarantee protection from hearing damage. However, personal monitors not only offer improved sound quality and convenience, but they also provide performers with an unprecedented level of control. Reducing stage volume also improves the listening experience for the audience, by minimizing feedback and interference with the house mix. As with most new technologies, an adjustment period is usually required, but few performers will ever return to floor monitors after using personal monitors.

For more information on hearing health for musicians, contact one of the following organizations.

House Ear Institute, Hotline: (213) 483-4431, Web site: www.hei.org
H.E.A.R., Hotline: (415) 409-3277, Web site: www.hearnet.com
Sensaphonics Hearing Conservation, 660 N. Milwaukee Avenue, Chicago, IL 60622
Toll Free: (877) 848-1714, Int'l: (312) 432-1714, Fax: (312) 432-1738
Web site: www.sensaphonics.com, E-mail: saveyourears@sensaphonics.com

References

Bob Ghent, "Healthy Hearing and Sound Reinforcement," *Mix*, Mar. 1994, pp. 150—162.
Susan R. Hubler, "The Only Ears You've Got," *Mix*, Oct. 1987, pp. 104–113.
G. Gudmundsen, "Occlusion Effect," *Etymotic Research Technical Bulletin*, Oct. 2000.
Thom Fiegle, "Notch Filter Allows for Best Monitor Mix Ever!" http://www.shure.com/ProAudio/TechLibrary/index.htm, Oct. 2005.

This page intentionally left blank

Chapter 38

Virtual Systems

by Ray Rayburn

38.1 The Design of Sound Systems 1439
  38.1.1 Analog Systems 1439
  38.1.2 The Introduction of Digital Devices in Analog Systems. 1440
38.2 Digitally Controlled Sound Systems 1440
  38.2.1 Digitally Controlled Analog Devices 1440
  38.2.2 Digitally Controlled Digital Audio Devices 1441
  38.2.3 Integration of Digital Products into Analog Systems 1442
    38.2.3.1 Dynamic Range 1442
    38.2.3.2 Level Matching 1442
    38.2.3.3 Level Matching Procedure 1442
    38.2.3.4 Minimization of Conversions 1442
    38.2.3.5 Synchronization 1443
    38.2.3.6 Multifunction Devices 1444
    38.2.3.7 Configurable Devices 1445
38.3 Virtual Sound Processors 1446
38.4 Virtual Sound Systems 1452
  38.4.1 Microphones 1452
  38.4.2 Loudspeakers. 1453
  38.4.3 Processing System. 1454
  38.4.4 Active Acoustics 1455
  38.4.5 Diagnostics 1455
  38.4.6 The Sound System of the Future 1455

This page intentionally left blank

38.1 The Design of Sound Systems

Sound systems are made of three primary components:

- Input transducers.
- Signal processing.
- Output transducers

38.1.1 Analog Systems

Transducers are devices that convert energy from one form into another.

The primary type of input transducer used in sound systems is the microphone. It converts the form of acoustic energy we call sound into electrical energy carrying the same information. Other common audio input transducers include the magnetic tape head, the optical sensor, the radio receiver, and the phonograph pickup cartridge. Tape recorders, floppy and hard drives use magnetic heads to transform analog or digital magnetic patterns on the magnetic media into electrical signals. Optical free-space links, optical fiber receivers, and CD and DVD players all use optical sensors to turn optical energy into electrical energy. Radio receivers turn carefully selected portions of radio frequency energy into electrical energy. Phonograph cartridges turn the mechanical motion of the grooves in a record into electrical energy.

Similarly, the most common type of output transducer used in sound systems is the loudspeaker. It converts electrical energy back into the form of acoustical energy we call sound. Other common output transducers include headphones, magnetic tape heads, lasers, radio transmitters, and record cutting heads. Headphones are specialized electrical to acoustic transducers, which are intended to produce sound for one person only. Tape recorders, floppy and hard drives use magnetic heads to transform electrical signals into magnetic patterns on the magnetic media. Optical free-space links, optical fiber transmitters, CDR, CDRW, DVD±RW, and BD recorders all use lasers to turn electrical energy into optical energy. Radio transmitters turn electrical signals into radio frequency energy. Phonograph cutting heads turn electrical energy into the mechanical motion of the grooves in a record.

In general, we can't just connect a microphone to a loudspeaker and have a usable sound system. While there are exceptions such as "sound powered" telephones, in almost all cases there needs to be something that falls under the general heading of signal processing to connect the two.

In its most simplified form this processing might only consist of amplification. In general, microphones have low electrical power output levels, while loudspeakers require more electrical input power in order to produce the desired acoustic output level. Thus the processing required is amplification.

The next most common form of audio signal processing is the level control. This is used to adjust the amount of amplification to match the requirements of the system at this moment.

Multiple inputs to the signal processing are often each equipped with their own level control, and the outputs of the level controls combined. This forms the most basic audio mixer.

Much of the rest of what is done in signal processing can be classified as processing that is intended to compensate for the limitations of the input and output transducers, the environment of the input and output transducers, and/or the humans using and experiencing the sound system. Such processing includes, among other things, equalization, dynamics processing, and signal delay.

Equalization includes shelving, parametric, graphic, and the subcategories of filtering, and crossovers among others. Common filters include high pass, low pass, and all pass. Crossovers are made of filters used to separate the audio into frequency bands.

Dynamics processing is different in that the parameters of the processing vary in ways that are dependent on the current and past signal. Dynamics processors include compressors, limiters, gates, expanders, automatic gain controls (AGC), duckers, and ambient level controlled devices.

Signal delays produce an output some amount of time (usually fixed) after the signal enters the device.

The biggest early breakthrough in sound systems was the development of analog electrical signal processing. No longer was the designer limited to some sort of mechanical or acoustic system. This was taken a large step forward with the development of vacuum tube-based electronic circuitry

Later, transistor circuitry allowed smaller product sizes and more complex processing to be done. The development of analog integrated circuits accelerated this trend.

Analog signal processing had its limitations, however. Certain types of processing such as signal delays and reverbs were very difficult to produce. Every time the signal was recorded or transmitted, quality was lost. Cascades of circuitry required to meet the ever more complex requirements of sound systems had reduced dynamic range when compared to the individual building blocks making them up.

38.1.2 The Introduction of Digital Devices in Analog Systems

These factors combined together spurred the application of digital signal processing (DSP) to audio.

DSP is the application of numerical processes to perform signal processing functions on signals that have been converted into sequences of numbers. The conversion of an analog signal into a sequence of numbers is done by an analog to digital converter or A/D. Similarly, the conversion of a sequence of numbers back into an analog signal is done by a digital to analog converter, or D/A.

Delays in the analog world almost always involved acoustic, mechanical, or magnetic systems. In other words, you had to use transducers to go from the electrical realm to some alternative form of energy and back, since it was very difficult to delay the signal enough to matter for audio systems while staying strictly in electrical form.

Early digital signal delays had very poor performance compared to today's digital products, but they were popular since the available analog delays were often of even worse audio quality, had very short maximum delay times, and often were not adjustable in delay time.

Digital signal delays offered much longer delay times, easy adjustability of the delay time, and often multiple outputs, Fig. 38-1.

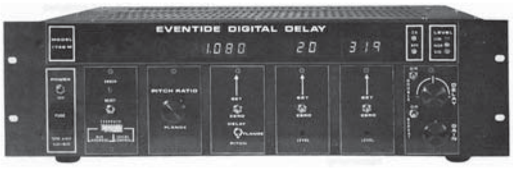


Figure 38-1. Introduced in 1974, the Eventide Clockworks Model 1745M modular digital audio delay was the first to use random access memory (RAM) instead of shift registers for storage of sound. Options included pitch changing and flanging.

Analog reverbs always required some sort of mechanical, magnetic, or acoustic system.

The first analog reverbs were simply isolated rooms built with very reflective surfaces, and equipped with a loudspeaker for inserting the sound, and one or more microphones for picking up the reverberated sound. Obviously, there are major drawbacks to this approach. The size and cost limited the application of this technique, as did the difficulty in adjusting the characteristics of the reverberation.

Another analog technique involved vibrating a large thin steel plate with sound, and picking up the vibrations at multiple locations on the surface of the plate. This had the advantage of smaller size, and adjustable reverberation by moving acoustic damping materials closer to or farther away from the plate.

Smaller yet analog reverbs were made using a gold foil instead of the steel plate, or by using springs that were driven at one end and had vibration pickups at their other end. The gold foil technique resulted in quite acceptable sound, but the spring-based systems were often of low cost and barely usable sound quality.

The first digital reverbs were very expensive, and of a size comparable to that of the gold foil or better spring systems, but had the advantage of much greater control over the reverberation characteristics than could be achieved with the analog systems. As the cost of digital circuitry has come down over the years, so has the price and size of DSP-based reverbs.

Analog recording and transmission of sound have always involved significant reduction in the sound quality as compared to the original sound. Each time the sound was rerecorded or retransmitted, the quality was further reduced.

Digital recording and transmission of sound offered a dramatic difference. While the conversion of analog signals into digital always involves some loss, as long as the signal was kept in digital form and not turned back into an analog signal, making additional copies or transmitting the signal did not impose any additional losses. Therefore, the generational losses associated with the analog systems we had been using were eliminated.

38.2 Digitally Controlled Sound Systems

38.2.1 Digitally Controlled Analog Devices

The physical controls of an audio device have always constituted a significant portion of the cost of the product. This was fine for devices such as mixing consoles where the operator needed instant access to all the controls. There have always been controls, however, that while necessary to the initial setup of the sound system, were best hidden from easy adjustment by the user. These controls were often placed behind security covers to reduce the chance of their adjustment by unauthorized users.

Crossovers, system equalizers, and delays are some of the more common examples of this class of device.

Until the introduction of the inexpensive graphically oriented personal computer, there was no practical way to avoid this cost, or to provide other than physical limitation of access to controls. Once graphical computers became commonplace and inexpensive, we saw the introduction of a new class of digitally controlled analog devices. These devices had greatly reduced numbers of physical controls, and in some cases had no physical controls. Instead the devices were adjusted by connecting them to a personal computer and running a control program on the computer. Once the "controls" were set using the program, the computer could be disconnected and the device would retain the settings indefinitely.

Since there were no or a limited subset of actual physical controls, there was no need for physical access restriction devices. Without both a computer and the appropriate control program, there was no way for the user to adjust the controls.

Presets, which could be recalled either from the front panel or remotely, now became possible. Different system configurations, which required multiple changes to controls, now became as simple as pressing a single button.

Computer control also allowed many more controls to be provided on a physically small product than would otherwise be practical. For example, a ⅓ -octave equalizer might have 27 bands, plus a number of presets, and yet could be packaged in a small box.

Remote control of the device allows the control point to be distant from the audio device itself. This opened the possibility of reducing the amount of audio cabling in a system and replacing it with inexpensive data cabling to the operator's control point. The data cabling is much more resistant to outside interference than audio cabling.

In the initial versions of such control systems, a different control program and physical connection from the computer to the audio device was required for each device that was to be so controlled. This was fine in smaller systems, but in larger installations where there might be many such digitally controllable devices, it quickly became cumbersome.

To address this limitation, Crown developed what they called the IQ system. It used a single control program together with a control network that connected many digitally controllable devices. Thus it provided a single virtual control surface on a computer screen, which allowed the adjustment and monitoring of multiple individual audio devices.

38.2.2 Digitally Controlled Digital Audio Devices

Early digital audio devices had physical controls that mimicked the controls of analog devices. As with digitally controlled analog devices, the advantages of remote control programs quickly became apparent, particularly for those devices with many controls.

Digital audio devices already were internally digitally controlled, so providing for remote control was an easy and relatively inexpensive step. Some such devices provided physical controls that communicated with the signal processor. Others provided only control programs that would run on a personal computer and connect via a data connection back to the device controlled.

As with the digitally controlled analog devices, most such control schemes required an individual data line from the control computer to each device controlled. Several manufacturers including TOA and BSS developed techniques to allow many of their devices to be controlled by a single data line. These schemes were limited to products of a single manufacturer. Work went on for many years under the auspices of the Audio Engineering Society to try to develop a universally applicable common control scheme, but the requirements were so diverse that a universal standard has yet to be achieved.

This was one of the factors leading to the rise of universal control systems from companies such as Crestron and AMX that have the ability to control and automate remote controllable equipment using any control protocol. For the first time such control systems allow the user to have a single control surface that operates all these systems with their diverse control protocols. These control systems control audio, video, lighting, security, and mechanical systems, allowing a degree of total system integration never before achieved.

Despite the success of these universal control systems, often the user just needs to control all his or her audio system components from a single interface. This has been a driving force behind the continued efforts to develop a common control protocol, or some other easy way to bring all these controls together for the user. Besides the work that the AES has done toward developing such a common protocol, control and monitoring protocols developed for other industries have been adapted for use with audio systems. Among these protocols are Simple Network Management Protocol, or SNMP, and Echelon LonWorks.

The desire for unified systems with reduced control interfaces has also been one reason for the popularity of integrated devices that combine the functions of many formerly discrete devices in a single unified product.

38.2.3 Integration of Digital Products into Analog Systems

38.2.3.1 Dynamic Range

The dynamic range of analog systems is characterized by a noise floor, nominal operating level, and maximum output level before a rated distortion is exceeded. The noise floor is usually constant and does not change with the audio signal. Distortion generally increases with increasing level. The increase in distortion as the maximum output level is approached may be gradual or sudden. Most professional analog audio equipment today has a maximum output level in the range from +18 to +28 dB relative to 0.775 V (dBu), and a nominal operating level in the 0 to +4 dBu range. While optimum operation requires matching of the maximum output levels so that all devices in the signal chain reach maximum output level at the same time, in practice many engineers do not bother to match maximum output levels, relying instead on the easier matching of nominal levels. The best signal levels to run through typical analog equipment are moderate levels, far away from the noise floor, but not too close to the maximum output level.

Digital equipment, on the other hand, has a different set of characteristics. Distortion decreases with increasing level, and reaches the minimum distortion point just before the maximum output level. At that maximum output level distortion rises very abruptly. The noise floor of digital equipment is often not constant. In some cases the noise is very signal dependent, and sounds to the ear much more like distortion than noise. These characteristics come together to suggest that the optimal signal levels would be those close to but just a little below the maximum level.

38.2.3.2 Level Matching

As we combine analog and digital equipment in the same system, the different characteristics of the two technologies suggest that for maximum performance and widest dynamic range we must align the maximum output levels.

Each device has its own dynamic range, but those ranges will have different characteristics. In all cases we want the audio signal to stay as far as possible from the noise floor.

In any system some device will have the smallest dynamic range, and thus set the ultimate limitation on the performance of the system.

The system as a whole may not perform as well as the worst performing component, unless care has been taken to assure that all devices reach their own maximum output level at the same time.

38.2.3.3 Level Matching Procedure

To match the maximum output levels, apply a midfrequency tone to the input of the first device in the system. Increase the applied level and/or the gain of the device until the maximum output level is reached as determined by the increase in distortion. This point may be determined by using a distortion meter, watching the waveform using an oscilloscope for the onset of clipping, or listening with a piezo tweeter connected to the output of the device.

This latter technique was developed by Pat Brown of Syn-Aud-Con. If a frequency in the range of 400 Hz is selected, the piezo tweeter can't reproduce it, and will remain silent. When the device under test exceeds its maximum output level, the resulting distortion will produce harmonics of the 400 Hz tone that fall in the range the piezo tweeter can reproduce, and it will sound off in a very noticeable way. The level is then reduced until the tweeter just falls silent, and maximum level has been determined. Rane has produced a commercial tester based on this concept called the Level Buddy.

Once the first device is at its maximum output level, the gain of the second device is adjusted to achieve its maximum output level. In some cases the input of the second device will be overloaded by the maximum output level of the first device and no adjustment of the gain of the second device will eliminate the distortion. In such a case, an attenuator must be used between the devices to drop off the level so the second device's input is not overloaded. One place where such an attenuator is often needed is at the input of a power amplifier. Many times professional power amplifiers have input overload points far lower than the maximum output levels of any of the common devices used to drive them.

Once the second device is at maximum output level, the process is repeated in turn for each subsequent device in the system.

Once all device interfaces have been optimized, the system is capable of maximum possible performance.

38.2.3.4 Minimization of Conversions

Up until now, digital components have been treated just like the analog components they have replaced in the system. This is not the optimal way, however, to integrate digital processing into a system.

All devices, analog or digital, impose quality limitations on the performance of the system. In a properly designed digital device, the major performance limitations are due to the conversions from analog into digital and back. Properly done DSP will not introduce significant distortions or other artifacts into the signal. Early analog to digital converters had significantly worse distortions than early digital to analog converters. With modern converters the pendulum has swung back the other way with analog to digital converters often having less distortion than digital to analog converters. In any case, the majority of the distortions and other quality degradations in a properly designed digital based audio signal processor are due to the converters.

This suggests that we should consider carefully how many A/D and D/A converters are used in our systems, with an eye to minimize the number of converters in any given signal path.

This requires a change in how we treat digital audio devices in our systems. No longer can we consider them to be interchangeable with traditional analog components. Instead we must use system design practices that will allow a reduction in the number of converters our audio must go through.

One powerful technique is to group all the digital devices together in only one part of the signal flow of our systems, and use digital interconnects between the devices instead of analog. While not all of our digital processors are available with digital interconnections, many of them are.

The most popular two channel consumer digital interconnect standard is known as SPDIF, while the most popular two channel professional interconnect standard is AES3. The European Broadcasting Union (EBU) adopted the AES3 standard with only one significant change. The EBU required transformer coupling, which was optional under AES3. As a result, this interconnect standard is often called AES/EBU. Many products are made with these interconnects, and converters are available to go from SPDIF to AES3 and from AES3 to SPDIF.

There are many interfaces that carry more than two channels. One that is popular in the home studio market is the ADAT interface, which carries eight channels. Most of the interfaces that originated in the home studio market are limited in the distance they can be run.

To address the need for greater distances and larger numbers of channels in professional applications, CobraNet was developed. It also differs from the other digital interfaces in that it allows point to multipoint connections instead of only point to point. This is due to it running on Ethernet, which is an industry standard computer networking protocol. Today there are several different digital interface systems sold that use some or all of the Ethernet Standard to transmit digital audio.

By grouping as many of our digital devices in one portion of the system as possible, and making all the interconnections between them in the digital domain, we have minimized the number of conversions our signal has gone through, and maximized the potential performance.

38.2.3.5 Synchronization

Digital audio consists of a series of consecutive numeric samples of the audio, each of which must be received in sequence. If the samples are not received in proper sequence, or if samples are lost or repeated, then the audio will be distorted.

In order for digital audio devices to interconnect digitally, both ends of each connection must run at the same sampling rate. If the source is running at even a very slightly faster rate than the receiver, sooner or later the source will output a sample that the receiver is not ready to receive yet. This will result in the sample being lost. Similarly, if the source is running at even a very slightly slower rate than the receiver, eventually the receiver will be looking for a sample before the source is ready to send it. This will result in a new false sample being inserted into the data stream.

In a simple chain of interconnected digital audio devices, it is possible for each device to look at the sampling rate of the incoming digital audio, and lock itself to that incoming rate. One problem with this system is that the sampling rate as recovered from the incoming digital audio is less than a perfect steady rate. It will have slight variations in its rate known as jitter. While there are techniques available to reduce this jitter, they add cost, and are never perfect. Each consecutive device in the chain will tend to increase this jitter. As a result, it is not recommended to cascade very many digital audio devices in this manner.

If a single digital audio device such as a mixer will be receiving digital audio from more than one source, then this simple scheme for synchronizing to the incoming digital audio signal breaks down, since there is more than one source. There are two ways to solve this problem.

One way is to use a sample rate converter (SRC) on each input to convert the incoming sample rate to the internal sample rate of the processor. Such a SRC will add cost to the input, and will in some subtle ways degrade the quality of the audio. Of course, there are different degrees of perfection available in SRCs at correspondingly different levels of complexity and cost. Some SRCs will only handle incoming digital audio that is at a precise and simple numeric ratio to the internal sample rate. Others will accept any incoming sample rate over a very wide range and convert it to the internal sampling rate.

This second sort of SRC is very useful when you must accept digital audio from multiple sources that have no common reference, and convert them all to a common internal sampling rate.

As implied above, the other way to handle inputs from multiple digital audio sources is to lock all the devices in the digital audio system to a common reference sample rate. In large systems this is the preferred solution, and the Audio Engineering Society has developed the AES11 Standard, which explains in detail how to properly implement such a system. Such a system can have excellent jitter performance since each device directly receives its sampling rate reference from a common source. Interconnections between the digital audio devices can be rearranged freely since we do not have to be concerned about synchronization and jitter changes as the signal flow is changed.

The only flaw in this scheme, is that some digital audio devices may not have a provision for accepting an external sampling rate reference. As a result, in many complex systems while there may be a master sample rate clock that most equipment is locked to, there often is still a need for samplthat can't lock to the master clock, or that operate at a different sample rate.

38.2.3.6 Multifunction Devices

Once we grouped most or all of our digital devices in a single subsection of our system, the next natural question is why not combine these multiple separate devices into a single product. Obviously, such a combined device greatly reduces or eliminates the need for the system designer to be concerned with synchronization issues, since the equipment designer has taken care of all the internal issues. Only if the system contains more than one digital audio device with digital interconnections does the issue of synchronization arise.

Some of the first examples of such combination products were digital mixers and loudspeaker processors.

Digital mixers were developed that combined not only the traditional mixing and equalization functions, but also often added reverb and dynamics processors inside the same device. Depending on the intended application, such digital mixers might also integrate automation systems, and remote control surfaces. Remote control surfaces allow the separation of the signal processing from the human operated controls. This might allow all the signal processing to remain on stage, for example, while only the control surface is placed at the operator's position, Fig. 38-2.

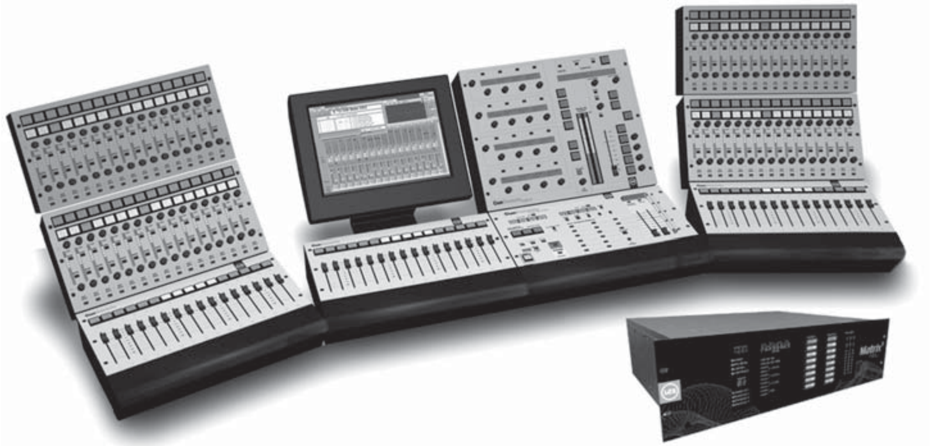


Figure 38-2. The CueConsole from Level Control Systems is a modular audio control surface. The size of the control surface has no direct relationship to the number of audio inputs and outputs controlled. The actual audio processing is performed in Matrix3 processors (lower right corner) located remotely from the control surface. These systems are very popular for Broadway- and Las Vegas-style shows since very large and powerful automated consoles can take up very little of the valuable space in the theater.

Loudspeaker processors are another common example of integrated digital subsystems. Such devices might include input-level adjustment, compression, signal delay, equalization, and crossover functions. Each crossover output might include further signal delay, equalization, level adjustment, and limiting. Often manufacturers provide standard settings for their loudspeakers using these processors, thus optimizing the performance of the loudspeaker to a degree not otherwise possible, while allowing one universal processor to be used for many different products in their line.

The limitation of such products is that their internal configuration is fixed, and therefore the possible applications are limited to those the manufacturer anticipated.

One solution is the one pioneered by Dave Harrison in his analog console designs. In an age when most recording consoles were custom built, and provided just the features and signal flow capabilities requested by the studio owner, David designed a console with so many features, and such flexible signal routing, that it could meet the needs of a very wide range of users. Any one user may not need any but a small subset of the available features. A few users might have requested additional features if they were having a custom console manufactured. Through innovative engineering, David was able to design this console in such a way that it could be mass produced for significantly less cost than the more limited custom consoles it replaced.

Applying this same concept to integrated digital devices led to devices designed with signal processing and routing capabilities well beyond the average user's requirements. This, of course, made such a device capable of application to more situations than a more limited device would have been.

38.2.3.7 Configurable Devices

The next significant advance in integrated digital signal processing was the user configurable device. In such a device, the basic configuration of the signal flow and routing remains constant, or the user can select from one of several different possible configurations. Next, the user can select the specific signal processing that takes place in each of the processing blocks in the selected configuration, within certain constraints.

This sort of device is fine for situations where the basic functions needed are limited, but some degree of customization to suit the job is required. The TOA Dacsys II was an early example of this sort of system, and was available in two in by two out and two in by four out versions, Fig. 38-3.

For example, a complex processor for a loudspeaker might have multiple inputs optimized for different types of audio inputs. There might be a speech input that is bandlimited to just the speech frequency range, equalized for speech intelligibility, and has moderate compression. There might be a background music input that has a wider frequency range, music-oriented equalization, and heavy compression. There might be a full range music input which has music equalization, and no compression.

The input processing chain for each will have a level control and a high pass filter for subsonic reduction or speech bandwidth reduction. The speech input chain might next have a low pass filter to reduce the high-frequency range. All three inputs will then have multiband parametric equalizers to tailor their frequency response. The speech and background music inputs would then have compressors for dynamic range control. The three input processing chains would end in a mixer that would combine them into a single mixed signal to drive the output processing chains.

Such a system might have three outputs, one for the low frequencies, one for the midfrequencies, and one for the high frequencies. The low-frequency processing chain will have a high-pass filter set to eliminate frequencies below the reproduction range of the woofer. Next, it would have a low-pass filter to set the crossover

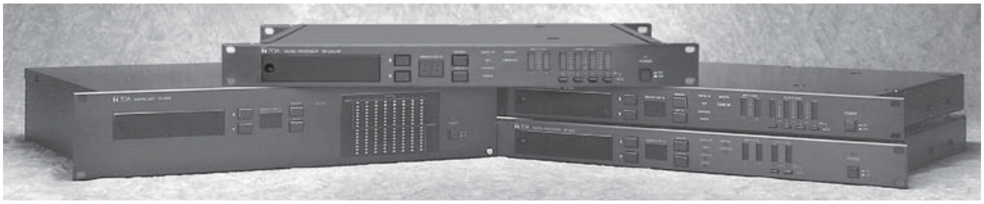


Figure 38-3. TOA Dacsys II digital audio processors (center and right). TOA's second generation digital audio processor (the SAORI was the first), it had a Windows-based control program, which allowed a limited amount of internal reconfiguration of the signal flow. On the lower left is a digitally controlled analog matrix mixer, which could be controlled with the same program.

frequency to the midrange speaker. This may often be followed by a multiband parametric equalizer used to smooth the response of the woofer. Lastly will come a limiter to keep the power amplifier from clipping and/or provide some degree of protection for the woofer. The midfrequency processing chain will have a high-pass filter to set the crossover frequency from the woofer, and a low-pass filter to set the crossover frequency to the high-frequency speaker. It might also have a multiband parametric equalizer and a limiter. The high-frequency processing chain will have a high-pass filter to set the crossover frequency from the midrange, and might have a low-pass filter to set the high-frequency limit. It may have a shelving equalizer to compensate for the high-frequency response of the driver. It will also have a multiband parametric equalizer and a limiter.

Some of these fixed configuration audio processors can allow quite complex systems to be built. For example the BSS ProSys ps-8810 provides eight inputs and ten outputs. Each input has filtering, delay, gating, AGC, compression, automatic mixing, more filtering, polarity, and muting available. This is followed by and eight in by ten out matrix mixer. Each output has delay, filtering, ambient level control, and limiter available. The combination of all these facilities allows quite complex systems to be built, Fig. 38-4.

38.3 Virtual Sound Processors

At times, however, even the most complex fixed configuration processor will not meet the needs of a project. In 1993, the Peavey MediaMatrix system introduced the concept of the virtual sound processor. It allowed the designers to choose from a wide variety of virtual audio processing devices, and wire them in any configuration they desired. Integrated digital sound systems could now be designed with the same flexibility of configuration formerly enjoyed in the analog world, and with much greater ease of wiring. Changes to the configuration could be rapidly made on screen and loaded into the processor at the click of a button. Complex systems, which would not have been possible using analog technology due to circuitry drift or cost, now became routine. Systems with as many as 256 inputs and 256 outputs, and 70,000 or more internal controls became practical, Fig. 38-5. More recently, BSS came out with the Soundweb digital audio processor with similar capabilities but a dif-

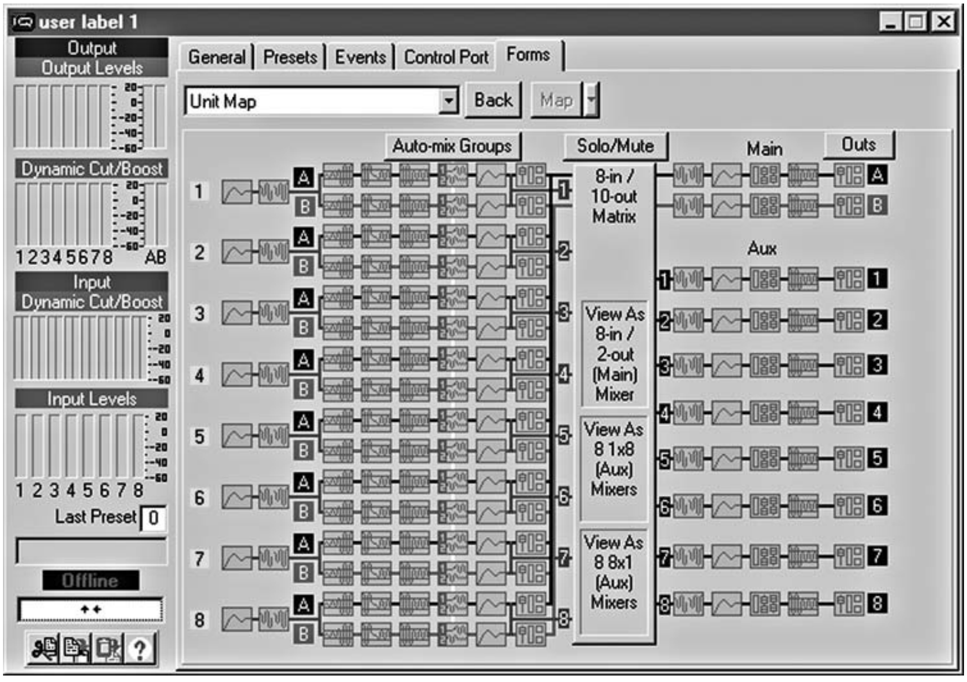
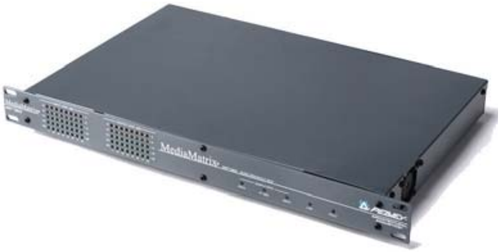

user label 1
Output
Output Levels
Dynamic Cut/Boost
12345678 AB
Input
Dynamic Cut/Boost
Input Levels
1 2 3 4 5 6 7 8
Last Preset 0
Offline
General
Presets
Events
Control Port
Forms
Unit Map
Back
Map
Auto-mix Groups
Solo/Mute
Main
Outs
8-in / 10-out Matrix
View As 8-in / 2-out (Main) Mixer
View As 8 1x8 (Aux) Mixers
View As 8 8x1 (Aux) Mixers
Aux
1
2
3
4
5
6
7
8
A
B

Figure 38-4. BSS control software for the ProSys ps-8810 showing the signal flow diagram. There is a limited ability to configure the function of the various signal processing blocks. This is a part of the IQ for Windows software package and can use CobraNet for digital I/O.

ferent physical architecture, Fig. 38-6A. MediaMatrix was based on a PC that supports digital audio processing cards inserted in it. What is now called Soundweb Original consists of a family of digital audio processor boxes that interconnect using category 5 UTP cables, hence the web part of the name. Both products provide similar functions in different ways.

A. Peavey MediaMatrix Frame 980nt. The first Virtual Sound Processor. It is a modular system based on a Windows/Intel PC with added Digital Processing Units (DPU's). From one to eight DPU's may be inserted as needed based on the amount of processing to be done.



B. Peavy MediaMatrix Break Out Box (BoB). Each BoB provides eight audio inputs and eight audio outputs for a MediaMatrix processor.

Figure 38-5. Peavey MediaMatrix mainframe system, the first virtual sound processor.

Biamp, BSS Audio, Electro-Voice, Innovative Electronic Designs, Level Control Systems, Peavey, QSC Audio Products, Symetrix, Yamaha, and others have come up with a variety of products that also give the user the ability to wire virtual devices together. These range from large systems similar to MediaMatrix or Soundweb, to a small module from QSC that provides processing for a single power amplifier, Fig. 38-6A–E.

Many of these signal processors provide multiple options for audio input and output. For example, MediaMatrix provides options for analog I/O, AES3 I/O, and a CobraNet interface.

In order for a virtual sound processor to replace all the analog processing used in a sound system, a wide variety of virtual devices must be available. MediaMatrix now provides nearly 700 standard audio processing and control logic virtual devices on its menu. It is also possible to build your own complex devices from simpler devices appearing on the menu or existing inside menu devices. Almost any audio processing device desired may be either found on the menu or built from components available.

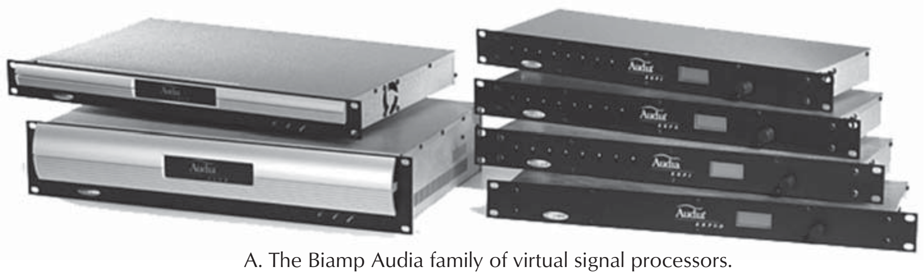
Systems are designed in a manner very similar to drawing a schematic on a CAD system. Virtual devices are taken from the menu and placed on a work surface. They have audio input nodes on the left, and audio output nodes on the right side of the device. Some systems have control input nodes on the top, and control output nodes on the bottom of the devices. Wires are drawn interconnecting the I/O nodes and the virtual devices, Fig. 38-7.

Any number of virtual devices may be used until the available DSP processing power is exhausted. All of the systems provide some means for displaying the amount of DSP used. Devices may be added to the schematic until 100% utilization is reached. Expandable systems such as MediaMatrix and Soundweb allow the addition of more cards or boxes to add additional processing power as needed. MediaMatrix also allows the selection of sampling rate. Slower sampling rates trade off reduced bandwidth for increased processing capability.

Since the schematic may be edited at any time, one major advantage of these systems is that changes may easily be made in the field to accommodate changed requirements or field conditions. Since it is rare that a system is 100% utilized, often the needed additional virtual devices, or wiring changes, may just be added. If the change exceeds the available DSP resources, often some other change may be made in a less critical area to reduce the required DSP resources. By contrast, in an analog system physical rewiring or the purchase of additional components would be required. Both of these add significant cost. Thus often the use of virtual sound processors results in significant savings in the total project cost, over and above the cost savings of the initial equipment purchase, and a more optimized finished system.

Generally, double-clicking on a virtual device will open it, allowing the internal controls to be seen, Fig. 38-8. Inside each device is a control panel with the controls and indicators needed by that device. Sometimes seldom used controls will be placed in sub-windows.

Selected controls from the devices may be copied and placed in control panels. This is done using the standard Windows copy and paste commands. The schematic may then be hidden, and the user only allowed access to the controls that the designer wishes, placed on the control panels. Multiple controls may be



A. The Biamp Audia family of virtual signal processors.

B. One of the BSS Audio London series of virtual signal processors.

C. Peavey's MediaMatrix NION N6 virtual signal processor.

D. QSC's BASIS series of virtual signal processors.

E. Yamaha's DME Satellite Series virtual signal processors.

Figure 38-6. Virtual sound processors by Biamp, BSS, Peavey, QSC, and Yamaha.

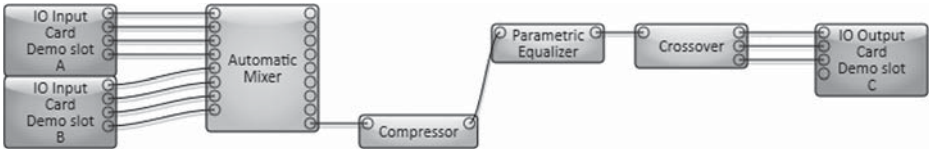
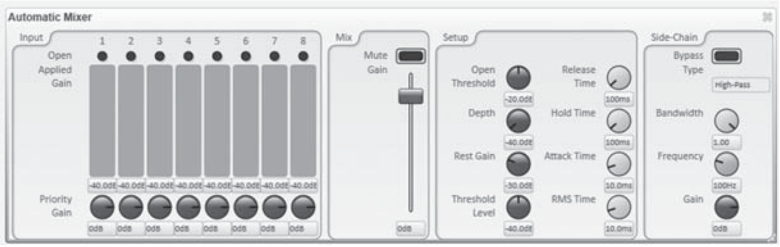

IO Input Card Demo slot A
IO Input Card Demo slot B
Automatic Mixer
Compressor
Parametric Equalizer
Crossover
IO Output Card Demo slot C

Figure 38-7. A simple example of a virtual sound processor schematic from QSC.

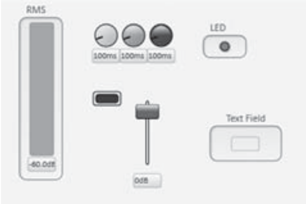
ganged, and the settings of many controls may be recalled using presets or subpresets. Presets recall the settings of all the controls in the system, while subpresets recall the settings of just a selected subset of the controls. Controls may also be edited to change their style, size, color, and orientation. This capability allows the designer to develop a very user-friendly interface. Often bitmaps may be inserted to serve as backgrounds.

Besides the virtual interface, some systems require physical interfaces. To support this requirement, most virtual sound processors provide remote control capability in addition to their virtual control surfaces. Some have a few front panel controls available, Fig. 38-9A. Many virtual sound processors provide control inputs to which external switches or level controls may be connected. Control outputs allow lamps and relays to be driven. Serial control interfaces using RS 232, RS 485, or MIDI are often available. Some processors also provide Ethernet interfaces. Others have dedicated programmable remote control panels. When remote control needs are extensive, but the user interface must be simple, touch screen operated control systems such as by AMX or Crestron are often used. These usually control the virtual audio processor by means of serial RS232, RS485, or Ethernet control lines, Figs. 38-9A– E.

Designing and using a virtual sound processor is similar to designing an analog system, except that you have the ability to more precisely optimize the system. The cost of each individual virtual device is very low, and you have the ability to wire precisely the configuration you need. Thus designs may be more efficient, and may also more exactly meet the system requirements.



A. Inside a typical Virtual Device from QSC. This shows the sort of controls and indicators found inside the Virtual Devices in Fig. 38-7.



B. A sampling of the control and indicator styles available in QSC. Controls inside Virtual Devices may be copied into control panels, and arranged into user friendly control screens.

Figure 38-8. Inside a typical virtual device from QSC. This shows the sort of controls and indicators found inside the virtual devices in Fig. 38-7.

Let's take our previous example of a loudspeaker processor with three inputs, each optimized for a different type of program, and three-way outputs, and design a virtual sound processor for this task using QSC. Other virtual sound processors could also be used for the same purpose although some details would be different.

The first step is to place the virtual devices for the audio inputs and outputs. Since we wish analog inputs and outputs, we will select analog I/O cards for inputs and outputs 1 through 4 from the Device menu, Fig. 38-10.

Input 1 will be our speech input. We will use a high pass filter, a 1 band parametric equalizer, a high-frequency shelving equalizer, a compressor, and a three input mixer. Input 1 will wire to the input of the high pass filter. The output of the high pass filter feeds the parametric equalizer, which feeds both the main input of the compressor, and the shelving equalizer. The output of the compressor wires to the first input of the mixer. The high pass filter will be adjusted to be a 125 Hz Butterworth 24 dB/oct filter. This band limits the input to the speech range, and prevents the entry of low-frequency noise. The parametric equalizer will be set for a bandwidth of two octaves, and 3 dB boost at 3 kHz. This provides a gentle emphasis of the speech intelligibility range. The compressor is left at its default settings of Soft Knee, 0 dB threshold, and a ratio of 2:1. The high-frequency shelving equalizer will be set to a frequency of 8 kHz, and 8 dB of boost. In combination with the compressor, this serves as a de-esser. By boosting the sibilance range at the input to the side chain of the compressor, those frequencies will be compressed more easily, and excessive hig-frequency sibilance will be controlled, fIG. 38-11.

Input 2 will be for the background music. We will use a high pass filter, two bands of parametric equalization, and a compressor. The high pass filter will be set to 80 Hz with a Q of 2. This produces an underdamped response with a bass boost just above the low-frequency roll-off. One band of the parametric equalizer is set to 1.5 kHz with a bandwidth of two octaves, and a cut of 5 dB, while the other is set to 8 kHz with a bandwidth of one octave, and a boost of 5 dB. The combination of the high pass filter and the parametric equalizer produces the desired background music response. The compressor is set to Soft Knee, –10 dB threshold, and a ratio of 4:1. This provides a more aggressive compression. The output of the compressor is wired to the second mixer input.

Input 3 is for full range music. It has a high pass filter, and a low-frequency shelving equalizer. The high pass filter is set to 30 Hz at 12 dB/oct, and the low-frequency EQ is set to +10 dB at 100 Hz. The output of the EQ is wired to the third mixer input.

The output of the mixer will drive a 6 band parametric equalizer for overall system EQ. Next comes a three-way 24 dB/oct crossover. The low-frequency output of the crossover is wired to a high pass filter with the *Q* adjusted so it optimally tunes and protects the woofer. Next comes a three band parametric EQ, five millisecond delay, and limiter. The side chain input of the limiter is wired directly from the output of the EQ bypassing the delay. This combination of a delay and limiter wired so that the main input of the limiter sees a

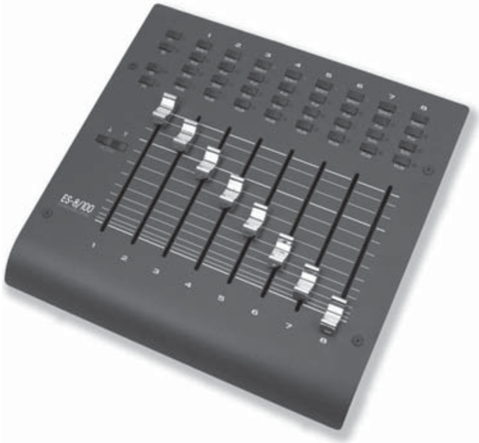
MediaMatrix
PEAVEY

A. Peavey MediaMatrix X-Frame88. This is an example of a Virtual Sound Processor which provides front panel controls which may be associated with internal controls in the virtual schematic.

Presets
Polarity
Mute
Main Menu
Left 0.0dB
Right 0.0dB
Level 0.0dB
Soundweb
BSS
Soundweb
BSS

B. BSS Audio Soundweb 9010 Remote Control. This is an example of a dedicated remote control panel that may control internal Soundweb controls. Six buttons, a rotary encoder, and a LCD display are provided.

C. BSS Audio Soundweb 9012 Wall Panel. This is an example of a simple remote control plate for a Virtual Sound Processor.



E. AMX NXD-CV17 touch screen control surface that can be used with virtual sound processors.

D. JL Cooper's ES-8/100 motorized fader package that can interface to virtual sound processors.

Figure 38-9. Examples of physical controls for virtual sound processors.

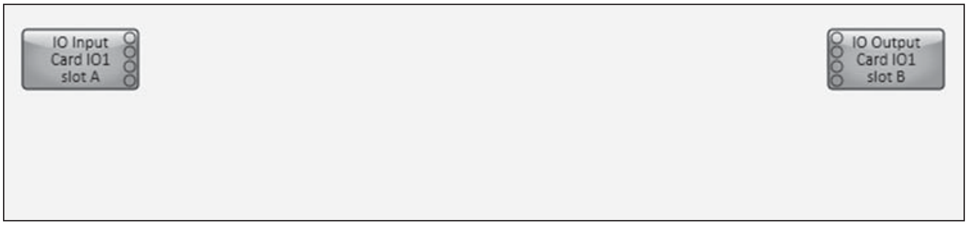

IO Input
Card IO1
slot A
IO Output
Card IO1
slot B

Figure 38-10. QSC software showing virtual devices for audio input and output.

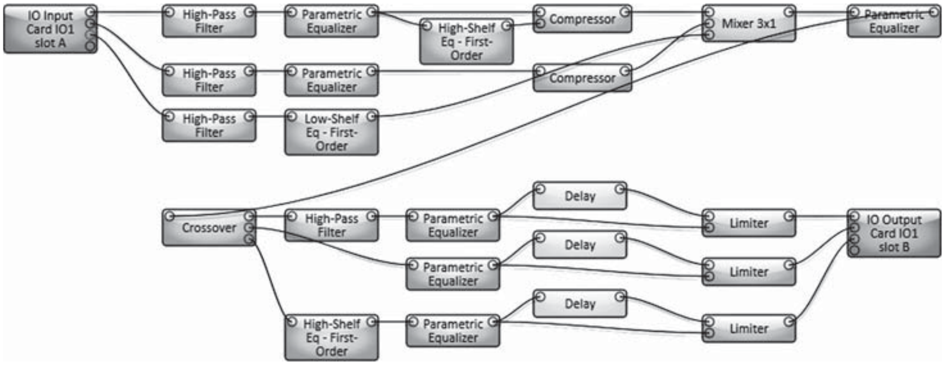


Figure 38-11. Completed crossover example schematic.

delayed signal, while there is no delay on the side chain input, is known as a *look-ahead limiter*. By setting the delay to about three times the attack time of the limiter, the limiter has time to react to an audio signal before that signal reaches the limiter. Since the limiter has an attack time of 1 ms, we will set the delay to 3 ms. A look-ahead limiter is able to accurately limit audio transients without the distortions inherent in ultrafast limiters.

The mid- and high-frequency outputs of the crossover are processed similarly. The midfrequency output just has the parametric EQ, delay, and limiter, while the high-frequency output has a shelving equalizer to compensate for the CD horn, parametric EQ, delay, and limiter, Fig. 38-11.

As you can see, while this circuit is relatively simple, using a virtual audio processor allowed us to optimize it in ways not possible using either commonly available analog components or a fixed configuration digital processor. This schematic utilizes about 3% of the resources in the small version of the QSC virtual processor. By way of comparison a similar schematic took 19% of the available DSP resources on a single MediaMatrix board. This shows the great improvement in DSP processing speed in the latest generation of virtual processors.

The larger and more complex the system, the greater the advantages of the virtual audio processor over previous technologies. Legislative chambers, stadiums, ballrooms, theme parks, and churches are among the facilities utilizing virtual audio processors.

One technique commonly used in legislative sound systems is called *mix-minus*. Often such systems will have a microphone and loudspeaker for each legislator. In order to prevent feedback, each loudspeaker receives a mix that does not contain its associated microphone signal. Signals from other nearby microphones are at a reduced level in the mix. The U.S. Senate sound system utilizes this technique. Since there are 100 senators each of whom has his or her own microphone and loudspeaker, plus leadership microphones and loudspeakers, there were over 100 microphones with over 100 associated loudspeakers, which would have required over 100 mixers each with over 100 inputs if it had been implemented with a straightforward matrix mixer. To reduce this complexity, the mix-minus technique was developed. It works on the concept that only a small number of microphone inputs need to be muted or reduced in level on any given output. A single large mixer is used to produce a mix of all the inputs called the sum. Each output mixer receives the sum and just those inputs that must be muted or reduced in level. The polarities of the direct inputs of the mixer are reversed, so that as their level is increased, they cancel out part or all of their audio from the sum at the output of the mixer. If a direct input is set to unity gain, it will perfectly subtract from the sum signal, thus eliminating that input from the mixer output. While this technique has been used in analog system designs, circuit stability restricted its practical use in larger systems. Digital systems add another potential complexity. As signals are processed and transferred between DSP processing chips, delays may be introduced. If the sum and direct input signals do not arrive at the mixer at exactly the same time, the direct signal will not properly cancel. Small amounts of signal delay on selected inputs may be required to assure that all the signals reach the input of any given mixer at the same time. Some virtual signal processors automatically provide such compensation, or provide it as an option. It is always possible to manually insert very small delays as required.

Today, virtually all larger sound reinforcement systems utilize some form of virtual sound processor. The advantages of more optimized system design, the ability to make easy changes, and reduced cost, have made this sort of processor the overwhelming favorite of consultants and contractors worldwide.

Systems installed today still use analog microphones and microphone preamps, and in some cases analog mixing consoles, the outputs of which are fed to the virtual sound processor. Likewise, the outputs of the virtual sound processors are usually connected in the analog domain to conventional power amplifiers, which are then wired to loudspeakers. Thus considerable portions of the total sound system remain outside the scope of the virtual sound processor. There is a better way, however, that was first used by the U.S. Senate sound system installed in 1994, and continued in the new system installed in 2006.

Each senator has his or her own small microphone equipped with a tiny Kevlar reinforced cord. Suitable microphones with direct digital outputs were not available. The cord is managed by a Servoreeler Systems servo controlled reeler located in the senator's desk, under the control of the virtual system. No slip rings are used, and the far end of the cord is directly connected to the preamp, also located in the desk. The analog gain of the preamp is also under the control of the virtual system. The output of the preamp drives an A/D converter, which is connected to a DSP processor, and also located in the desk. The initial audio and control processing is done in the desk. Audio, control signals, and power are carried on a single cable between the desk and the central processor portion of the virtual sound system. The central processor performs all the mix-minus processing, and many of the auxiliary functions. Outputs from the central processor go back over the same cable to the desk, where further processing is done, still under the control of the virtual system, and the power amplifiers and speakers are driven.

What is special about this system is that not only is the central processing done in a virtual sound processor, but the processing associated with the microphones and loudspeakers is also part of the virtual system. The entire system consisting of redundant central processors, over 100 desk units, custom operator's console, and several displays, is all part of a single integrated virtual sound system. All of the DSP processing, including both central processors, and the over 100 remote processors in the desks, is loaded with their operating code and controlled, from a single common virtual sound system program. Even the microphone reelers, which are a servo controlled electromechanical subsystem, and the analog microphone preamps, are under the control of the virtual sound system. There are no unnecessary A/D and D/A conversions, and the longest analog interconnection is the microphone cable. Sound is converted from analog into digital at the end of the microphone cable, and remains in the digital domain until it is in the loudspeaker enclosure. The U.S. Senate sound systems, both the original of 1994 and the updated 2006 system, can be considered to be prototypes for the virtual sound systems of the future.

38.4 Virtual Sound Systems

38.4.1 Microphones

The virtual sound system of the future will be programmed and controlled through a single unified user interface program. It will have no analog interconnections.

Microphones will have a direct digital output, and receive power and control signals through the microphone cable. The Audio Engineering Society Standards Committee has issued the AES-42-2006 Standard defining a digital interface for microphones. The digital audio transmission scheme used is based on the AES3 Standard, but adds digital phantom power, microphone control, and synchronization features. The microphones can phase-lock their internal sampling clocks to that of the equipment to which they are connected. The first microphones meeting this standard contain conventional analog microphone elements with conversion into the digital domain done inside of the microphone body. In the future, we may see microphones that produce digital signals directly out of the microphone element. In either case, these new smart digital microphones can be controlled by the virtual system to which they are connected. Some of these microphones will even allow their directional patterns to be changed, and in some cases to be steered toward the sound sources under the control of the virtual sound system. By dynamically adjusting the pickup pattern and direction of each of the microphones, the sound system may adaptively optimize its performance.

Microphone arrays will enhance the control of the directional patterns and aiming of microphones. Microphone arrays consist of from three to hundreds of microphone elements whose outputs are processed to produce one or more virtual microphones with controllable directional patterns and orientation. They will have the ability to produce narrow pickup patterns if so desired, which can be aimed dynamically at the desired sound source. If the sound source moves, the pickup pattern can also move to follow the sound source. Because of this capability, array microphones will be capable of picking up intelligible sound from a greater distance than traditional microphones. This will allow sound systems to be designed without visible microphones, even in difficult acoustic environments.

Since the outputs of the individual microphone elements in an array are processed into the single virtual microphone output by DSP processing, adding more DSP will allow additional virtual microphones to be produced from the same array. Each virtual microphone can be aimed in a different direction, with its own directional pattern. We already see the beginnings of this in some of the microphone systems that provide 5.1 surround outputs for recording from a single compact microphone array. In a speech reinforcement system, each virtual microphone could track an individual talker. Talkers would be identified by their unique individual voiceprint. When there are multiple microphone arrays in the room, each talker's individual virtual microphone can automatically be formed using the optimum nearby array. As each talker moves around the room, his personal virtual microphone will always be formed using a nearby array, and will move from array to array as he moves around the room.

Since each virtual microphone will stay with its assigned talker, the output of each microphone may be individually optimized for the person to which it is assigned. When logging of the activities in the room is required, if desired each person could be recorded on her own individual track. Where speech to text conversion is utilized, having a separate virtual microphone for each talker is a significant advantage. Speech to text conversion is much easier when the system can learn the voice of a single individual. By providing outputs to the speech to text system that only contains the voice of a single individual, accuracy is greatly improved.

Microphone arrays will also have the ability to selectively reject sounds coming from certain sources. During system setup, the virtual microphone processors will be taught the location of the system loudspeakers, and of any significant noise sources in the room. This will allow them to keep a null in the directional pattern always aimed in those directions. As a result, the chances of feedback and the pickup of noise will be significantly reduced.

It will also be possible to define regions in 3D space from which speech will not be amplified. In legislative systems, for example, it is extremely important to make sure side conversations are never amplified. By defining an area slightly back from the desks as a privacy zone, the legislators will be able to lean back and have a private conversation with their aides even if they forget to turn their microphones off.

Current voice tracking microphone arrays are limited in their bandwidth, add significant signal latency, and are costly. These factors have made them unattractive for sound reinforcement applications. However, improvements in processing algorithms, coupled with the dramatic reductions in the cost of DSP processing power we have seen each year, will soon bring this technology to a host of new applications including sound reinforcement.

38.4.2 Loudspeakers

Many loudspeakers today are powered with integrated power amplifiers and crossovers. Some loudspeakers have expanded on this concept by directly accepting digital audio and control signals. They contain DSP processing, which, integrated with the loudspeaker system design, allow much improved loudspeaker performance and protection. Modern DSP-based line array loudspeakers have steerable directional patterns, and in some cases can produce multiple acoustic output beams from the same loudspeaker. They may even send back an audio sample of their acoustic output for confidence monitoring.

As with microphone arrays, DS-based loudspeaker arrays allow sound to be steered to where it is needed, and kept from where it is not wanted. Dynamically controlled loudspeaker arrays will allow the loudspeaker coverage to change as room and system conditions change. Loudspeaker arrays may be produced as lines or flat panels, which mount flush with the walls, ceilings, and other architectural room elements. No longer is it necessary for loudspeakers to be aimed in the direction we wish the sound to go. For example, it is quite feasible to mount a flat panel loudspeaker array in a convenient location on the sidewall of the room, and direct the sound downwards and back into the audience area. Loudspeaker coverage patterns and directions may be changed under the control of the virtual system for different uses of the facility. This is a tremendous advantage over the older technolog, which required either multiple sets of speakers, or physically changing the loudspeaker aiming for different applications.

A single loudspeaker array may be used to simultaneously produce multiple sound coverage patterns, each of which may be driven by its own independent sound source if so desired. One application of this technique would allow greatly enlarging the area in a room where accurate multichannel reproduction could be heard. Those located towards the edges of the room could now receive properly balanced sound from all loudspeakers in the room, even though they were much closer to some of the loudspeakers than to others, thus preserving the spatial reproduction. In another application, the same loudspeaker could aim direct sound at the audience, while simultaneously aiming ambient effects toward other portions of the room.

Control feedback from the virtual sound system will allow automatic modification of the loudspeaker coverage pattern as environmental conditions change. Such changes might include audience size and location, ambient noise, temperature, and wind speed and direction. Integration of DSP processing will also allow other useful functions to be moved into the loudspeaker cabinet. These will include source signal selection and mixing, delay, equalization, compression, limiting and driver protection, and ambient level compensation. The programming and control of the DSP processing will be over the same connection that brings the digital audio to the loudspeaker. This will allow the integration of all loudspeaker functions as part of the common virtual sound system.

38.4.3 Processing System

Those portions of the audio processing that are not contained either in the microphones or the loudspeakers will be contained in the central processing system. This may be either a single processor, or a networked array of processors. In either case there will be a single user interface for programming and controlling the entire system.

Control and monitoring of the virtual sound system may occur from many locations concurrently. The system will be controllable from PCs running either dedicated control software, or even standard Web browsers. For situations where control via a mouse is not acceptable, touch screen controllers will be available. Where physical controls are desired, a variety of standard, modular, control panel elements will be available. These will allow implementation of physical controls as simple as a wall mounted volume control, or as complex as a large mixing console.

Virtual sound processors have evolved substantially since the first products of this type were introduced in the early 90s. As the processing power available in these products has grown so have the capabilities.

Sound systems exist in a real-world environment, which also contains many other elements with which the sound system operation must be integrated. The most advanced of today's virtual sound processors contain powerful control logic subsystems to ease this integration. High-speed control connections allow exchange of data with external room and building control systems.

QSC Audio recently introduced a new audio processing product suite that has advanced the reliability, sophistication, and capabilities of virtual audio processing products. The QSC offering incorporates many functions that previously were available only in distinctly separate products. These include advanced virtual devices such as FIR filters, feedback suppressions, and ambient level sensing. It also greatly reduces the amount of time needed to compile, as well as incorporates a very low, and fixed, latency between all inputs and outputs. The QSC product allows the designer to easily create a fully redundant system, answering much of the concern that was initially generated by the use of digital systems for all of a facility's audio signal processing and control.

One very significant advantage of the most advanced virtual sound processing systems is the ease with which it is possible to make the various processing subsections interact with each other. For example, an automatic microphone mixer can be thought of as multiple-level meters and gain blocks, where the signal level at the various inputs is used to adjust the instantaneous gain of the various gain block. Such automatic microphone mixers exist in analog, digital, and virtual form. However, that sort of interaction can be expanded greatly to the system level in a virtual sound processor. For example, each microphone input processing chain might contain an AGC. The maximum possible gain an individual AGC can insert while still keeping the entire sound system stable will depend in part on the amount of gain or loss the AGCs for every other microphone are applying. In a virtual system it is possible to let each AGC know what the other AGCs are doing, and based on that information modify its behavior.

There are devices on the market that dynamically insert notch filters to keep a sound system from going into feedback. They do this by monitoring the onset of feedback and very quickly applying the corrective filters. This means the system must slightly start to ring before correction can be applied. A virtual sound system, by contrast, can monitor all the factors that impact system stability and insert corrective notch filters selectively in only the signal path required, and do so before the system starts ringing.

A virtual sound system can be programmed to know which are the most critical microphones and loudspeaker zones, and if trade-offs must be made to get optimum performance, can optimize the most important inputs and outputs. For example, if there is a person who must be heard, and that person is speaking in a very soft tone of voice and as a result the gain can't be gotten high enough, the virtual sound system can bring the gain up higher in the most critical loudspeaker zones while not increasing the gain everywhere, thus keeping the overall system stable.

A virtual sound processor may have many thousands of controls that need to be adjusted during a systems initial setup. Today's advanced virtual processors contain control tools that allow the system commissioning engineer a much simplified interface for adjusting those controls. This greatly reduces the time required and the chances for error in setup.

In short a well-designed virtual sound system can apply all the little tweaks to the system's controls that a very skilled operator would have applied if he or she could respond to conditions in a split second and adjust hundreds of controls at once.

38.4.4 Active Acoustics

The virtual sound system may also be used to modify the acoustic environment.

The reverberation time and reflection patterns of the space may be dynamically varied at any time to meet the needs of the program material. This requires that the physical acoustics of the space be at the low end of the desired reverberation range. The virtual sound system will add the initial reflections from the proper spatial directions, and the enveloping reverberant tail, to produce the desired acoustic environment. The ability to change the acoustics on an almost instantaneous basis allows each portion of a program to be performed in its optimum acoustics. For example, spoken portions of the program may only utilize a few supportive reflections. At the other extreme, choral or organ music may have a very long reverberation time. This technology may also be used to simulate the acoustic environment of the room in outdoor performance venues.

Environmental noise, particularly that of a low-frequency and/or repetitive nature, may be actively canceled by the virtual sound system. As the cost of DSP processing comes down, and the power handling of transducers goes up, this technology will become more attractive in comparison to traditional noise control and isolation methods. Vibration and low- frequency sounds are the most difficult and costly to isolate using traditional passive methods. High displacement isolation mounts together with large amounts of mass are often required for good low-frequency performance. At higher frequencies often far less expensive techniques and materials are effective. By comparison, active noise and vibration control techniques are most effective at the lowest frequencies, but find it increasingly difficult to obtain satisfactory performance over large areas at higher frequencies. Therefore, including active noise control techniques in a virtual sound system to control low-frequency noises may prove beneficial in reducing the total project cost.

38.4.5 Diagnostics

The virtual sound system will monitor its own operation, and the environment in which it operates. The entire signal path will be monitored for failures. Depending on the level of system design, the operator may just be notified, or redundant equipment may be automatically utilized to assure uninterrupted operation. Most systems will utilize multiple microphones and loudspeakers. In itself, this provides a significant degree of redundancy. If the coverage pattern of the microphones or loudspeakers is controllable, then the virtual system can compensate for any given failure of a microphone or loudspeaker. Redundancy may also be designed into the interconnections and processing subsystems of the virtual sound system. With careful design, systems with few or no single points of failure can be built.

Environmental conditions that will impact the long term health of the system, such as temperature and airflow, will be monitored and trends logged. The performance of the microphones and loudspeakers in the system will be monitored and recorded to spot degradation of performance before it becomes audible. The acoustic environment will also be monitored to spot changes that might impact on the subjective performance of the sound system. System health reports will be automatically generated and sent to the system operator, installer, and designer when any of the parameters monitored are outside of expected tolerances. This capability will result in much more consistent performance over the life of the system, and will extend that life for years.

38.4.6 The Sound System of the Future

When all these techniques are combined, the virtual sound system of the future will have better performance, be more invisible to the user, be easier to operate, and have a longer life than any current system.

This page intentionally left blank

Chapter 39

Digital Audio Interfacing and Networking

by Ray Rayburn

39.1 Background 1461
  39.1.1 Synchronous Connections 1461
  39.1.2 Asynchronous Connections 1461
  39.1.3 Isochronous Connections 1462
  39.1.4 AES5 1462
  39.1.5 Digital Audio Interconnections 1462
  39.1.6 AES11 1463
39.2 AES3 1463
  39.2.1 Subframe Format 1464
  39.2.2 Frame Format 1464
  39.2.3 Channel Coding 1465
  39.2.4 Preambles 1465
  39.2.5 Channel Status Format 1466
39.3 AES3 Implementation 1466
  39.3.1 AES3 Transmitters 1466
  39.3.2 AES3 Receivers 1466
  39.3.3 Electrical Interface 1466
  39.3.4 Line Drivers 1471
  39.3.5 Jitter 1471
  39.3.6 Line Receivers 1471
  39.3.7 AES3 Connectors 1472
39.4 AES-3id 1473
  39.4.1 Line Driver 1473
  39.4.2 Interconnect System 1473
  39.4.3 Line Receiver 1473
39.5 AES42 (AES3-MIC) 1474
  39.5.1 Synchronization 1474
  39.5.2 Microphone ID and Status Flags 1474
  39.5.3 Remote Control 1474
  39.5.4 Simple Instruction Set 1475
  39.5.5 Direct Commands 1475
  39.5.6 Remote Control Pulse Structure 1477
  39.5.7 Synchronization 1477
  39.5.8 Microphone Identification and Status Flags 1478
    39.5.8.1 Organization 1478
  39.5.9 XLD Connector for AES3-MIC Applications 1481

39.6 IEC 60958 Second Edition 1482
  39.6.1 Electrical and Optical Interface 1483
    39.6.1.1 Unbalanced Line 1484
    39.6.1.2 Optical Connection 1484
39.7 AES10 (MADI) 1484
  39.7.1 NRZI Encoding 1485
  39.7.2 Sample Frequencies and Rates 1485
  39.7.3 Synchronization 1486
  39.7.4 Electrical Characteristics 1486
  39.7.5 Optical Interface 1486
39.8 Soundweb 1486
39.9 Nexus 1487
39.10 IEEE 1394 (FireWire) 1487
39.11 Ethernet 1487
  39.11.1 Ethernet History 1488
  39.11.2 Ethernet Packet Format 1488
  39.11.3 Network Diameter 1489
  39.11.4 Ethernet Varieties 1489
  39.11.5 Ethernet Topology 1490
  39.11.6 Ethernet Equipment 1490
  39.11.7 Ethernet Connection Negotiation 1494
  39.11.8 Managed Switches 1495
  39.11.9 Virtual Local Area Network (VLAN) 1496
  39.11.10 Quality of Service (QoS) 1496
  39.11.11 Routing 1496
  39.11.12 Fault Tolerance 1496
    39.11.12.1 Trunking 1496
    39.11.12.2 Spanning Tree 1496
    39.11.12.3 Meshing 1497
    39.11.12.4 Rapid Spanning Tree 1497
  39.11.13 Core Switching 1498
  39.11.14 Ethernet Wiring 1498
    39.11.14.1 UTP Cable Grades 1498
    39.11.14.2 Distance Limitations 1499
  39.11.15 Connectors 1499
    39.11.15.1 Pairing, Color Codes, and, Terminations 1500
  39.11.16 Fiber Optic Cable 1502
    39.11.16.1 Fiber Optic Connectors 1502
    39.11.16.2 Cabling and Network Performance 1503
  39.11.17 Cable Installation 1503
    39.11.17.1 Cable Ties and UTP 1503
    39.11.17.2 Pull Force and Bend Radius 1503
    39.11.17.3 Cat 5 1503
    39.11.17.4 Fiber Optic Cable 1503
    39.11.17.5 Cable Testing 1504
39.12 CobraNet® 1504
  39.12.1 CobraNet® Terminology 1504
  39.12.2 Protocol 1505
  39.12.3 Beat Packet 1505
  39.12.4 Isochronous Data Packet 1505

39.12.5 Reservation Packet 1505
39.12.6 Timing and Performance 1505
39.12.7 Bundle Identification 1506
39.12.8 Multicast Bundles 1506
39.12.9 Unicast Bundles 1506
39.12.10 Private Bundles 1507
39.12.11 Bundle Assignments 1507
39.12.12 CobraCAD® 1509
39.12.13 CobraNet® Hardware 1509
39.13 Aviom 1509
39.14 EtherSound 1510
39.15 Dante 1511
39.16 QSC 1511
References 1511

This page intentionally left blank

39.1 Background

In most cases it is preferred to interface digital audio devices in the digital domain, instead of using analog interconnections. This is because every time audio is transformed from analog to digital, or digital to analog, there are inevitable quality losses. While analog interfacing is simple and well understood, there are few cases in which it would be desirable to interface two digital audio devices in the analog domain. If the digital audio devices are not provided with digital audio interfaces, for example, analog interfacing will be required. Such an analog interface, however, will result in subtle changes in the digital audio from one side of the interface to the other. The exact sequence of numbers that make up the digital audio will not be reproduced at the far side of an analog interface.

The numbering system commonly used in digital audio is called binary. Each of the digits (called bits) in the binary numbering system can be either a 1 or a 0. If two binary numbers are identical, then all their bits will match.

Digital audio interfaces have the potential to allow bit accurate transfer of the digital audio from one digital audio device to another, thus insuring no changes in the sequence of numbers that makes up the digital audio, and therefore potentially perfect accuracy. In order for this potential to be realized both digital audio devices must be synchronized.

Digital audio consists of a series of consecutive numeric samples of the audio, each of which must be received in sequence. If the samples are not received in proper sequence, or if samples are lost or repeated, then the audio will be distorted.

In order for digital audio devices to interconnect digitally, both ends of each connection must run at the same sampling rate. If the source is running at even a (very slightly faster) rate than the receiver, sooner or later the source will output a sample that the receiver is not ready to receive yet. This will result in the sample being lost. Similarly, if the source is running at even a (very slightly slower rate) than the receiver, eventually the receiver will be looking for a sample before the source is ready to send it. This will result in a new false sample being inserted into the data stream.

39.1.1 Synchronous Connections

The most straightforward way to carry digital audio is over a synchronous connection. In such a scheme, the data is transmitted at the exact same rate it is created, in other words, at the sample rate. When additional data is sent along with the audio in a synchronous system, it is added to the audio data and the whole package of information is transmitted at the audio sampling rate. Such systems send information in fixed-size groups of data, and introduce very little signal delay of latency. In a synchronous system the audio data words are sent and received at the audio sampling rate and both ends of the system must be locked to the same master sampling rate clock, Fig. 39-1. AES3 and IEC 90958 are examples of synchronous digital audio interconnection schemes.

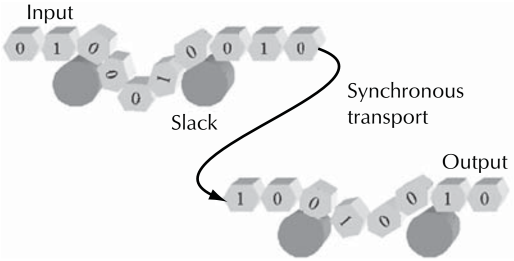


Figure 39-1. Synchronous connection between input and output.

39.1.2 Asynchronous Connections

An asynchronous system is in many ways the exact opposite of a synchronous system. Information is not sent at any particular time. The size of a given packet of information may vary. The time that it takes to get a given piece of information across an asynchronous connection may well be indeterminate. There is no common master clock that both ends of the connection refer to.

Examples of asynchronous transmission abound. When you mail a letter, it may contain just a short note on a single page, or it might contain the manuscript for a book. You put your letter in the mailbox (the outbound buffer) in the expectation that it will be picked up sometime later that day. The letter will pass through many different stages of transmission and storage along the way to its destination. You might know that the average delivery time is 3 days, however, in some cases the delivery might happen in 2 days, and in others it might be 6 days. You can be (almost) certain the letter will reach its destination eventually, but the exact delivery time can't be known.

Other examples of asynchronous transmission include the Internet, and most common computer interfaces and networks including RS-232 serial, and Ethernet networking.

RealAudio and Windows Media Audio (WMA) are two common schemes for providing a synchronous audio connection over the asynchronous Internet. While there can be a large receive buffer of as much as 5 to 10 seconds, you will still experience dropouts in the audio due to network conditions on the Internet inserting delay on some audio packets that exceed the receive buffer delay. Generally the audio gets through OK, but every once in a while it drops out.

39.1.3 Isochronous Connections

Isochronous connections share properties of both the synchronous and asynchronous systems, and bridge the gap between the two. Information is not sent at a constant rate locked to a master clock. It provides a maximum delivery time for information that is only rarely if ever exceeded. By using buffers at each end, it can carry information such as audio where the delivery of audio words in proper sequence, and with a known and constant delay, is essential. In a properly designed isochronous system latency can be very low providing near real time operation, and reliability can be very high, Fig. 39-2.

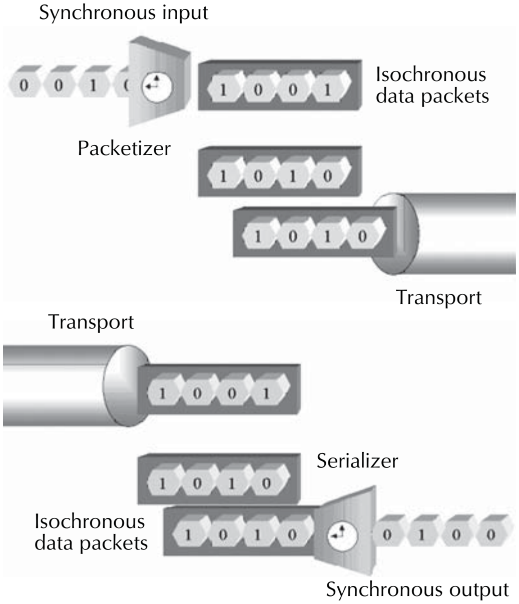


Figure 39-2. An isochronous system.

Examples of isochronous systems include ATM that is commonly used for transmitting telephone calls and computer data, and CobraNet® audio networking. FireWire (IEEE-1394) is a networking scheme that combines both isochronous and asynchronous elements.

39.1.4 AES5

AES5 standardizes on a primary sampling frequency for professional audio use of 48 kHz ±10 parts per million (ppm). It also allows 44.1 kHz to be used when compatibility with consumer equipment is required. For broadcast and transmission-related applications where a 15 kHz bandwidth is acceptable it allows a sampling frequency of 32 kHz to be used. For applications where a wider than 20 kHz bandwidth is desired, or a relaxed slope of the antialiasing filter is preferred, a sampling rate of 96 kHz ±10 ppm may be used.

Higher and in some cases much higher sampling rates are in use internally in digital audio equipment. When such a higher rate appears on an external digital audio interface, the AES recommends the rate be a multiple of a factor of two of one of the approved sampling rates above.

AES5 discourages the use of other sampling rates, although others are in use.

The above information is based on AES5-2003. It is always advisable to obtain the latest revision of the standard.

39.1.5 Digital Audio Interconnections

In a simple chain of interconnected digital audio devices, it is possible for each device to look at the sampling rate of the incoming digital audio, and lock itself to that incoming rate. One problem with this system is that the sampling rate as recovered from the incoming digital audio is less than a perfect steady rate. It will have slight variations in its rate known as jitter. While there are techniques available to reduce this jitter, they add cost, and are never perfect. Each consecutive device in the chain will tend to increase this jitter. If the jitter gets too high, the receiving device may not correctly interpret the digital audio signals, and bit accuracy will be lost. Worse, the performance of analog to digital and digital to analog convertors is very dependent on a precise and steady clock. Even very small amounts of jitter can significantly degrade the performance of convertors. As a result, it is not recommended to cascade very many digital audio devices in this manner.

If a single digital audio device such as a mixer will be receiving digital audio from more than one source, then this simple scheme for synchronizing to the incoming digital audio signal breaks down, since it is only possible to synchronize to a single source at a time. There are two ways to solve this problem.

One way is to use a sample rate converter" (SRC) on each input to convert the incoming sample rate to the internal sample rate of the digital audio device. Such a SRC will add cost to the input, and will in some subtle ways degrade the quality of the audio. The accuracy will be better than with an analog interfacing, but the digital audio will not be transferred with bit accuracy. Of course, there are different degrees of perfection available in SRCs at correspondingly different levels of complexity and cost. Some SRCs will only handle incoming digital audio that is at a precise and simple numeric ratio to the internal sample rate. Others will accept any incoming sample rate over a very wide range and convert it to the internal sampling rate.

This second sort of SRC is very useful when you must accept digital audio from multiple sources that have no common reference, and convert them all to a common internal sampling rate.

As implied above, the other way to handle inputs from multiple digital audio sources is to lock all the devices in the digital audio system to a single common reference clock rate. In large systems this is the preferred solution, and the Audio Engineering Society has developed the AES11 Standard that explains in detail how to properly implement such a system. Such a system can have excellent jitter performance since each device directly receives its sampling rate reference from a common source. Interconnections between the digital audio devices can be rearranged freely since we do not have to be concerned about synchronization and jitter changes as the signal flow is changed.

39.1.6 AES11

AES11 defines a digital audio reference signal (DARS) that is merely an accurate AES3 signal used as the common reference clock for a facility. The DARS may contain audio signals, but is not required to do so.

There are three basic modes of operation defined in AES11: use of a DARS, use of the embedded clock in the AES3 signal, and use of a common master video reference clock from which a DARS is derived. Use of a DARS is considered normal studio practice. As mentioned above cascading AES3 signals through devices without a DARS can lead to increased jitter.

The only flaw in this scheme is that some digital audio devices may not have provisions for accepting an external sampling rate reference. As a result, in many complex systems, while there may be a master sample rate clock that most equipment is locked to, there often is still a need for sample rate convertors to accept the output of those devices that can't lock to the master clock. AES11 acknowledges this limitation.

AES11 specifies two grades of DARS, grade 1 and grade 2. A DARS that has as its primary purpose studio synchronization should be identified in byte 4 bits 0–1 of the AES3 channel status. More details are given on this below.

A grade 1 DARS is the highest quality and is intended for use in synchronizing either a multiroom studio complex or a single room. It requires a long term stability of ±1 ppm. Devices producing a grade 1 DARS are only expected to themselves lock to signals of grade 1 quality. Devices that are only expected to lock to grade 1 signals are required to lock to signals over a range of ±2 ppm.

A grade 2 DARS is intended for use in synchronizing only within a single room where the added expense of a grade 1 solution can't be justified. It requires a long term stability of ±10 ppm as specified in AES5. Devices expected to lock to grade 2 signals are required to lock to signals over a range of ±50 ppm.

The above information is based on AES11-2003. It is always advisable to obtain the latest revision of the standard.

39.2 AES3

The Audio Engineering Society titled their AES3 Standard "Serial transmission format for two-channel linearly represented digital audio data." Let's break that title apart as our first step in examining the AES3 Standard.

This standard sends the data in serial form. In other words it sends the information to be transmitted as a sequence of bits down a single transmission path, as opposed to sending each bit down a separate transmission path. Each bit of data making up a single sample of the audio is sent in sequence starting with the least significant bit on up to the most significant bit. The least significant bit is the bit that defines the smallest change in the audio level, while the most significant bit is the one that defines the largest change in the audio level.

AES3 normally is used to transmit two channels of audio data down a single transmission path. The data for channel one of a given audio sample period is sent first, followed by the data for channel two of the same sample. This sequence is then repeated for the next sample period.

Most professional digital audio today is linearly represented digital audio data. This is also sometimes called pulse code modulation, or PCM. In such a scheme for numerically representing audio, each time sample of audio is represented by a number indicating its place in a range of equal-sized amplitude steps. If eight bits were being used to represent the audio level, there would be 2^8 or 256 equal-sized amplitude steps between the smallest level that could be represented and the largest. The smallest amplitude change that could be represented is exactly the same at the low-level portion of the range as at the highest-level portion. This is important to understand since not all digital audio uses such a linear representation. For example, often telephone calls are encoded using nonlinear techniques to maximize the speech quality transmitted using a limited number of bits. In professional audio we generally use larger numbers of bits, usually in the range of 16 to 24, that allow excellent performance with linear representation. Linear representation makes it easier to build high-quality converters and signal processing algorithms.

The bits that make up an audio sample word are represented in two's complement form, and range from the least significant bit (LSB) that represents the smallest possible amplitude change, to the most significant bit (MSB) that represents the polarity of the signal.

AES3 adds a considerable amount of structure around the basic sequence of bits described above in order to allow clock recovery from the received signal, provide a robust signal that is easily transmitted through paths of limited bandwidth, and provide for additional signaling and data transmission down the same path.

Each of the two audio channels that can be carried by an AES3 signal is formatted into a sequence of two subframes, numbered 1 and 2, each of which follows the following format.

The following information is based on AES3-2003. It is always advisable to obtain the latest revision of the standard.

39.2.1 Subframe Format

First, additional bits are added before and after the digital audio to make a subframe of exactly 32 bits in length. The bits are transmitted in the sequence shown from left to right, Fig. 39-3.

If the audio data contains 20 or fewer bits, subframe format B is used. If the audio data contains 21 to 24 bits, subframe format A is used. In either case if data containing less than 20 or 24 bits is used, extra zeros are added to the LSB to bring the total number of bits to 20 or 24. Since the data is in two's complement form, it is important that the MSB, representing the signal polarity, always be located in bit 27.

The preamble is used to indicate if the audio to follow is channel one or two, and to indicate the start of a block of 192 frames.

If 20 or less audio bits are carried, then AES3 allows 4 bits of other data to be carried by the AUX bits.

The validity bit is zero if it is permissible to convert the audio bits into analog, and one if the conversion should not be done. Neither state should be considered a default state.

The user bit may be used in any way for any purpose. A few possible formats for using this bit are specified by the standard. Use of one of these formats is indicated by the data in byte one, bits 4 through 7 of the channel status information. If the user bit is not used it defaults to zero.

The channel status bit carries information about the audio signal in the same subframe, in accordance with a scheme that will be described later.

The parity bit is added to the end of each subframe and is selected so the subframe contains an even number of ones and an even number of zeros. This is called even parity. It allows a simple form of error checking on the received signal.

39.2.2 Frame Format

A subframe from channel two follows a subframe from channel one. The pair of subframes in sequence is called a frame.

In normal use frames are transmitted at exactly the sampling rate.

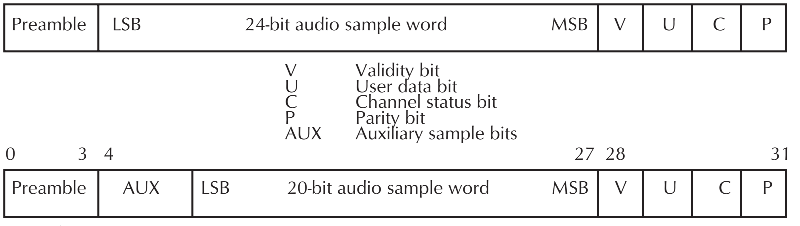


Figure 39-3. AES3 subframe format. Note that the first bit is called bit 0.

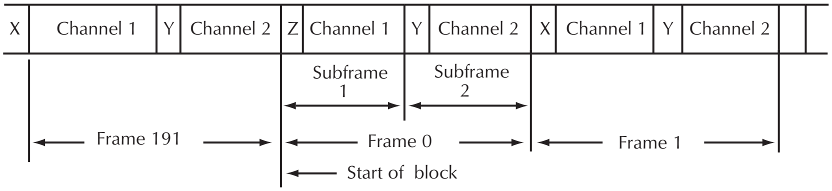


Figure 39-4. AES3 frame format. Note that the subframes are numbered 1 and 2, but frames are numbered starting with frame 0.

Again the data is transmitted in the sequence shown from left to right.

The parts shown as X, Y, and Z above represent the three versions of the preamble portion of each subframe. When version Z is used, it indicates the start of a block of 192 frames. When version X or Z is used, it indicates that the channel data to follow is from channel one. When version Y is used, it indicates that the channel data to follow is from channel two.

Blocks are used to organize the transmission of channel status data, Fig. 39-4.

39.2.3 Channel Coding

AES3 needs to be able to be transmitted through transformers. Transformers can't pass direct current (dc). Ordinary binary data can stay at 1 bit level for any arbitrary length of time, and thus by its nature can contain a dc component. Therefore a coding scheme is needed that eliminates this possibility.

We must also be able to recover the sampling rate clock from the AES3 signal itself. It was desired not to have to rely on a separate connection to carry the sampling rate clock. Since ordinary binary can stay at a given bit level for any arbitrary length of time, it is not possible to extract the clock from such a signal.

It was also desired to make AES3 insensitive to polarity reversals in the transmission media.

To meet these three requirements, all of the data except the preambles is coded using a technique called biphase-mark.

The binary data shown in the source coding portion of the diagram above has the sequence 100110.

The clock marks shown are at twice the bit rate of the binary source coding, and specify a time called the unit interval (UI), Fig. 39-5.

The channel coded data sequence has a transition at every boundary between bits of the original source coding, whether or not the source coding has such a transition. This allows extraction of the original clock rate from the received signal since there always is a transition at every source bit boundary.

If the source coding data is a one, the channel coding will insert a transition in the middle of the source coding bit time. If the source coding data is a zero, the channel coding will not insert any additional transition.

The combination of these channel coding characteristics provides the desired features. There is no dc component, so the signal may be transmitted through transformers. The sampling rate clock may be extracted from the signal. The signal is insensitive to polarity reversals since the source data state is carried by the presence or absence of an additional signal transition rather than the coded data state itself.

39.2.4 Preambles

The single portion of the subframe that is not encoded using biphase-mark coding is the preamble. In fact the preambles are deliberately designed to violate the biphase-mark rules. This is done to allow easy identification of the preamble and to avoid any possibility that some data pattern could by chance duplicate a preamble.

This also allows the receiver to identify the preamble and synchronize itself to the incoming audio within one sample period. This makes for a robust reliable transmission scheme.

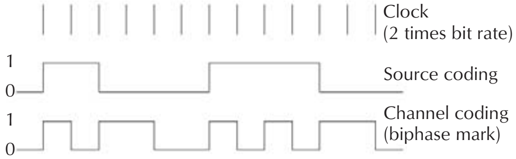


Figure 39-5. AES3 channel coding. The time between clock pulses is called the unit interval (UI).

As mentioned in the Frame Format section above, there are three different possible preambles. Each preamble is sent at a clock rate equal to twice the bit rate of the source coding. Thus the eight states of each preamble are sent in 4 bit time slots at the beginning of each subframe.

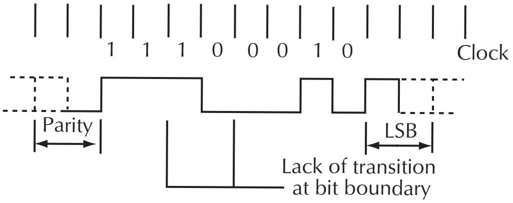


Figure 39-6. AES3 Preamble X (11100010). The time between clock pulses is called the unit interval (UI).

The state of the beginning of the preamble must always be opposite that of the second state of the parity bit that ends the subframe before it.

Preceding State	0	1	
	Channel Coding		
"X"	11100010	00011101	Subframe 1
"Y"	11100100	00011011	Subframe 2
"Z"	11101000	00010111	Subframe 1 and block start

You will note that the two versions of each preamble are simply polarity reversed versions of each other.

In practice, due to the nature of the positive parity used for the bit before the preamble, and the biphase coding, only one version of each preamble will ever be transmitted. However, to preserve the insensitivity to polarity inversions, AES3 receivers must be able to accept either version of each preamble.

Like biphase-mark coding, the preambles are dc free and allow for clock recovery while differing from valid biphase-mark coding at least twice.

The clock rate shown above is at twice the source bit rate. Note that the second state of the parity bit is always zero, and therefore the preamble will always start with a transition from zero to one. Also note that in this preamble, as in all possible preambles, there are at least two places where there is no transition at a bit boundary thus violating the rules for biphase-mark coding and providing positive identification of the preamble.

39.2.5 Channel Status Format

Each audio channel has its own channel status bit. The data carried by that bit is associated with its own audio channel. There is no requirement that the data for each channel be identical, although it could be, Fig. 39-7.

The sequence of 192 channel status bits in a given block is treated as 24 bytes of data, as shown in Table 39-1.

The sequence of channel status bits for each channel starts in the frame with Preamble Z.

39.3 AES3 Implementation

39.3.1 AES3 Transmitters

As a minimum, an AES3 transmitter must encode and transmit the audio words, validity bit, user bit, parity bit, the three preambles, and a minimum version of the channel status.

The minimum version of channel status will have byte 0 bit 0 set to one to specify this is a "professional use of channel status block" and all the other bytes set to their default values.

Some AES3 receiving devices might have problems with such a minimum version of channel status for two reasons. First, many receivers expect to see a properly encoded CRC in byte 23, and will therefore show a CRC error when receiving the default 0's instead of a CRC. Second, some receivers might expect to see the sampling frequency in byte 0 bits 6–7, and not have provision for manual override or auto set of the sampling frequency.

Even if some addition information is included in the channel status beyond what is listed as a minimum above, unless all the information considered standard below is included, the interface must still only be called a minimum implementation of AES3.

A standard implementation will include everything specified as minimum above plus will encode and transmit all the information in bytes 0, 1, 2, and 23 of the channel status.

An enhanced implementation provides additional capabilities beyond the standard implementation.

All transmitters must be documented as to which of the channel status capabilities they support.

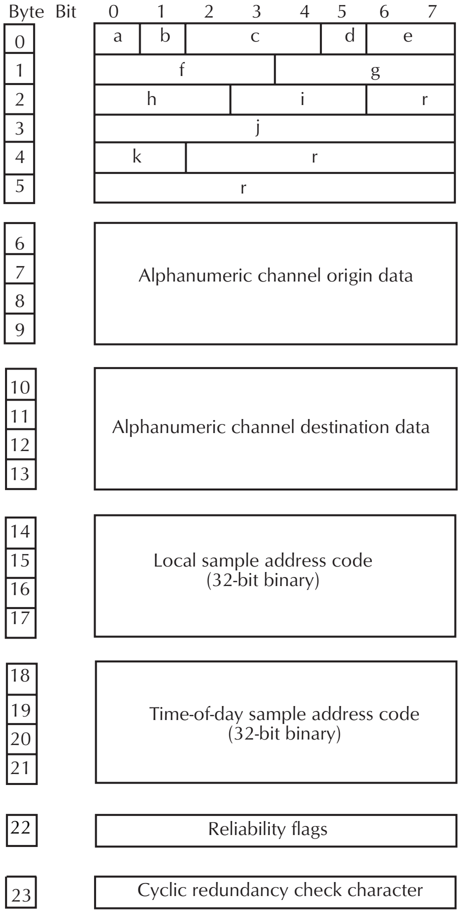
39.3.2 AES3 Receivers

All receivers must document the level of implementation provided and the actions that will be taken by the receiving device based on the information received.

39.3.3 Electrical Interface

AES3 uses a balanced 110 Ω electrical interface based on the International Telegraph and Telephone Consultative Committee (CCITT) Recommendation V.11. It is designed for use at distances up to "a few hundred meters."

If improved performance beyond that of CCITT V.11 is desired, it is suggested but not required that the circuit in Fig. 39-8 be used.


Byte
Bit
0 1 2 3 4 5 6 7
0
1
2
3
4
5
a
b
c
d
e
f
g
h
i
r
j
k
r
r
6
7
8
9
Alphanumeric channel origin data
10
11
12
13
Alphanumeric channel destination data
14
15
16
17
Local sample address code (32-bit binary)
18
19
20
21
Time-of-day sample address code (32-bit binary)
22
Reliability flags
23
Cyclic redundancy check character

a. Use of channel status channel
b. Audio/nonaudio use
c. Audio signal emphasis
d. Locking of source sample frequency
e. Sampling frequency
f . Channel mode
g. User bit management
h. Use of auxiliary sample bits
i. Source word length and source encoding history
j. Future multichannel function description
k. Digital audio reference signal
r. Reserved

Figure 39-7. AES3 Channel Status Data Format. Note both the bits and bytes are numbered starting with 0.

Series capacitors C_2 and C_3 block external dc from flowing through the transformers. This protects the transformers from damage or performance degradation if dc is applied to them. The AES42 (AES3-MIC) Digital Interface for Microphones Standard calls for digital microphones to be powered by 10 *Vdc* digital phantom power that is a variation on the phantom power scheme used for analog microphones. This provides another excellent reason to provide dc blocking on all AES3 inputs and outputs, transformer based or not, to prevent damage if such a phantom power scheme were to be applied. Structured wiring using Cat5 or high rated cable and RJ45 connectors is now a permitted alternative interconnect scheme for AES3 signal. These interconnects are also used for Ethernet, which may have power over Ethernet (PoE) applied. They are also used for plain old telephone service (POTS), which will have 48 V battery and 90 V ring signals. If structured cabling is used for AES3, consideration must be given to the survivability of the line driver and receiver circuits if accidentally interconnected to PoE or POTS lines.

Transformers will make possible higher rejection of common mode interfering signals, electromagnetic interference (EMI), and grounding problems than common active circuits. The European Broadcasting Union (EBU) in its version of this standard (EBU Tech. 3250-E) requires the use of transformers. This is the major difference between the standards. It is common to see the AES3 Standard referred to as the AES/EBU Standard even though that is not strictly correct since AES3 makes the transformers optional, while the EBU requires them.

Table 39-1. Channel Status Data Format Details

Byte 0		
Bit 0	0	Contents of the channel status block conform to IEC 60958-3 "consumer use" Standard. Ignore the rest of this table. (See Note 1.)
	1	Contents of the channel status block as to the AES3 "professional use" Standard.
Bit 1	0	Audio words consist of linear PCM samples.
	1	Audio words consist of something other than linear PCM samples.
Bits 2–4	Encoded Audio Signal Emphasis	
Bit	2 3 4	
State	0 0 0	No emphasis indicated. Receiver defaults to no emphasis but may be manually overridden.
	1 0 0	No emphasis used. Receiver may not be manually overridden.
	1 1 0	50/15 µs emphasis used. Receiver may not be manually overridden.

Table 39-1. Channel Status Data Format Details (Continued)

	1 1 1	International Telegraph and Telephone Consultative Committee (CCITT) J.17 emphasis (with 6.5 dB insertion loss at 800 Hz). Receiver may not be manually overridden.
	All other possible states of bits 2–4 are reserved and are not to be used unless defined by the AES in the future.	
Bit 5	0	Lock not indicated. This is the default.
	1	Source sampling frequency unlocked.
Bits 6–7	Encoded Sampling Frequency (See Notes 2, 3 and 4.)	
Bit	6 7	
State	0 0	No sampling frequency indicated. This is the default. Receiver defaults to 48 kHz, and automatic sampling rate determination or manual override is enabled.
	0 1	48 kHz sampling rate. Automatic sampling rate determination or manual overrides are disabled.
	1 0	44.1 kHz sampling rate. Automatic sampling rate determination or manual overrides are disabled.
	1 1	32 kHz sampling rate. Automatic sampling rate determination or manual overrides are disabled.

Note 1. Other than the use of the Channel Status block of information, the rest of the data format is identical between the AES3 "professional use" Standard and the IEC 60958-3 "consumer use" Standard. The electrical format is different, however. For these reasons it should never be assumed that a "consumer use" receiver would function correctly with a "professional use" transmitter, or vice versa.

Note 2. It is not a requirement that the sampling frequency used be indicated by these bits, nor is the use of one of the sampling frequencies that can be indicated by these bits. If the transmitter does not support sampling frequency indication, the sampling frequency is unknown, or the sampling frequency is not one that can be indicated by these bits, then the bits should be set to 0 0. Bits 3–6 of Byte 4 of the Channel Status may indicate other possible sampling rates.

Note 3. If a sampling rate is indicated, it may be modified by the status of Bit 7 of Byte 4 of the Channel Status.

Note 4. If Bits 0–3 of Byte 1 indicate single channel double sampling frequency mode, then the sampling frequency indicated by Bits 6 and 7 of Byte 0 is doubled.

Byte 1

Bits 0–3	Encoded Channel Mode	
Bit	0 1 2 3	
State	0 0 0 0	No mode indicated. Receiver defaults to two channel mode but may be manually overridden.
	0 0 0 1	Two channel mode used. Receiver may not be manually overridden.
	0 0 1 0	Single channel (monophonic) mode used. Receiver may not be manually overridden.
	0 0 1 1	Primary/Secondary (subframe 1 is primary) mode used. Receiver may not be manually overridden.
	0 1 0 0	Stereophonic (subframe 1 is left) mode used. Receiver may not be manually overridden.

Table 39-1. Channel Status Data Format Details (Continued)

	0 1 0 1	Reserved for user defined applications.
	0 1 1 0	Reserved for user defined applications.
	0 1 1 1	Single channel double sampling frequency mode. Subframes 1 and 2 contain successive samples of the same signal. Sampling frequency is double the frame rate and double the rate indicated in Byte 0 (if a rate is indicated there), but not double the rate indicated in Byte 4 (if a rate is indicated there). Receiver may not be manually overridden. Byte 3 may indicate channel number.
	1 0 0 0	Single channel double sampling frequency mode—left stereo channel. Subframes 1 and 2 contain successive samples of the same signal. Sampling frequency is double the frame rate and double the rate indicated in Byte 0 (if a rate is indicated there), but not double the rate indicated in Byte 4 (if a rate is indicated there). Receiver may not be manually overridden.
	1 0 0 1	Single channel double sampling frequency mode—right stereo channel. Subframes 1 and 2 contain successive samples of the same signal. Sampling frequency is double the frame rate and double the rate indicated in Byte 0 (if a rate is indicated there), but not double the rate indicated in Byte 4 (if a rate is indicated there). Receiver may not be manually overridden.
	1 1 1 1	Multichannel mode. Byte 3 indicates the channel numbers.
	All other possible states of bits 0–3 are reserved and are not to be used unless defined by the AES in the future.	
Bits 4–7	Encoded User Bits Management	
Bit	4 5 6 7	
State	0 0 0 0	No user information indicated—default.
	0 0 0 1	192 bit block of user data, starting with the Preamble "Z."
	0 0 1 0	Data to the AES18 Standard.
	0 0 1 1	User defined.
	0 1 0 0	Data to the IEC 60958-3 Standard.
	All other possible states of bits 4–7 are reserved and are not to be used unless defined by the AES in the future.	

Byte 2

Bits 0–2	Encoded use of auxiliary sample bits.	
Bit	0 1 2	
State	0 0 0	Maximum 20 bit audio words, auxiliary sample bits usage not defined, default.
	0 0 1	Maximum 24 bit audio words, auxiliary sample bits used for audio.
	0 1 0	Maximum 20 bit audio words, auxiliary sample bits used for a coordination signal per Annex A of AES3.
	0 1 1	User defined applications.
	All other possible states of bits 0–2 are reserved and are not to be used unless defined by the AES in the future.	

Table 39-1. Channel Status Data Format Details (Continued)

Bits 3–5	Encoded audio word length (see notes 1, 2, 3, and 4).		
Bit	3 4 5	Audio word length if bits 0–2 indicate maximum 20 bit length.	Audio word length if bits 0–2 indicate maximum 24 bit length.
State	0 0 0	Length not indicated, default	Length not indicated, default
	0 0 1	19 bits	23 bits
	0 1 0	18 bits	22 bits
	0 1 1	17 bits	21 bits
	1 0 0	16 bits	20 bits
	1 0 1	20 bits	24 bits
	All other possible states of bits 3 – 5 are reserved and are not to be used unless defined by the AES in the future.		
Bits 6–7	Alignment level indication.		
Bit	6 7		
State	0 0	Not indicated.	
	0 1	SMPTE RP155 (alignment level is 20 dB below maximum level).	
	1 0	EBU R68 (alignment level is 18.06 dB below maximum level).	
	1 1	Reserved for future use.	

Note 1. If the default state or bits 3–5 is indicated, the receiver should default to 20 or 24 bits as specified by bits 0–2, but allow manual override or auto set.

Note 2. If other than the default state of bits 3–5 is indicated, the receiver should not allow manual override or auto set.

Note 3. No matter which audio word length is indicated, the MSB representing the signal polarity is always bit 27 of the subframe.

Note 4. Knowledge of the actual encoded audio word length can be used to allow the receiving device to properly re-dither the audio to a different word length if so required.

Byte 3

Bit 7	Defines the meaning of bits 0–6.	
State	0	Undefined multichannel mode, default.
	1	Defined multichannel modes.
Bits 0–6	Channel number if bit 7 is 0. Channel number is value of bits 0–6 (bit 0 is LSB) plus 1.	
Bits 4–6	Multichannel mode if bit 7 is 1.	
Bit	4 5 6	
State	0 0 0	Multichannel mode 0. Bits 0–3 specify the channel.
	1 0 0	Multichannel mode 1. Bits 0–3 specify the channel.
	0 1 0	Multichannel mode 2. Bits 0–3 specify the channel.
	1 1 0	Multichannel mode 3. Bits 0–3 specify the channel.
	1 1 1	User defined multichannel mode. Bits 0–3 specify the channel.
Bits 0–3	Channel number if bit 7 is 1. Channel number is value of bits 0–3 (bit 0 is LSB) plus 1.	

Note 1. It is intended that the defined multichannel modes will identify mappings between channel numbers and function. Standardized mappings have yet to be defined.

Table 39-1. Channel Status Data Format Details (Continued)

Note 2. Some equipment may only consider the channel status data carried in one of the two subframes. Therefore if both subframes specify the same channel number, subframe 2 has a channel number one above channel 1 unless single channel double sampling frequency mode is in use.

Note 3. If bit 7 is 1, bits 0–3 correspond to the consumer mode channel status specified in IEC 60958-3. Consumer mode channel A is equivalent to channel 2, and Consumer mode channel B to channel 3 and so on.

Byte 4

Bits 0–1	Digital audio reference signal to the AES11 Standard.	
Bit	0 1	
State	0 0	This is not a reference signal, default.
	0 1	Grade 1 reference signal.
	1 0	Grade 2 reference signal.
	1 1	Reserved for future use.
Bit 2	Reserved.	
Bits 3–6	Sampling frequency.	
Bit	3 4 5 6	
State	0 0 0 0	No frequency indicated, default.
	1 0 0 0	24 kHz.
	0 1 0 0	96 kHz.
	1 1 0 0	192 kHz.
	0 0 1 0	Reserved.
	1 0 1 0	Reserved.
	0 1 1 0	Reserved.
	1 1 1 0	Reserved.
	0 0 0 1	Reserved for vectoring.
	1 0 0 1	22.05 kHz.
	0 1 0 1	88.2 kHz.
	1 1 0 1	176.4 kHz.
	0 0 1 1	Reserved.
	1 0 1 1	Reserved.
	0 1 1 1	Reserved.
	1 1 1 1	User defined.
Bit 7	Sampling frequency scaling flag.	
	0	No scaling, default.
	1	Multiply sampling frequency indicated in byte 0 bits 6–7, or byte 4 bits 3–6, by 1/1.001.

Note 1. The sampling frequency as indicated in byte 4 is independent of the channel mode as indicated in byte 1.

Note 2. There is no requirement to use a particular sampling frequency, nor to use a sampling frequency that can be indicated in bytes 0 or 4. If the transmitter does not support indication of sampling frequency, the frequency is unknown, or the sampling frequency is not one that can be indicated in this byte, then bits 3–6 should be set to "0 0 0 0."

Note 3. It is intended to assign sampling frequencies in the future to the currently reserved states of byte 4 bits 3– 6 (except 0 0 0 1) such that if the rates are related to 44.1 kHz bit 6 will be set, and if they are related to 48 kHz bit 6 will be cleared. Do not use these reserved states unless defined in the future by the AES.

Byte 5

Table 39-1. Channel Status Data Format Details (Continued)

Bits 0–7	Reserved. Set to 0 unless defined in the future by the AES.
Bytes 6–9	
	Alphanumeric channel origin data. Byte 6 contains the first character.
Bits 0–7 (each byte)	7 bit International Organization for Standardization (ISO) 646, American Standard Code for Information Interchange (ASCII), data. No parity bit is used. Bit 7 is always 0. Transmit LSBs first. Nonprintable characters (codes 01 to 1F hex and 7F hex) must not be used. Default is all 0’s (code 00 hex or ASCII null).
Bytes 10–13	
	Alphanumeric channel destination data. Byte 10 contains the first character.
(each byte)	7 bit ISO 646 ASCII data. No parity bit is used. Bit 7 is always 0. Transmit LSBs first. Nonprintable characters (codes 01 to 1F hex and 7F hex) must not be used. Default is all 0’s (code 00 hex or ASCII null).
Bytes 14–17	
	Local sample address code sent as 32 bit binary with LSBs first. Value is of the first sample in this block.
Bits 0–7	
(each byte)	Transmit LSBs first. Default is all 0’s.
Note 1. This serves the same function as an index counter on a recorder.	
Bytes 18–21	
	Time of day sample address code sent as 32 bit binary with LSBs first. Value is of the first sample in this block.
Bits 0–7	
(each byte)	Transmit LSBs first. Default is all 0’s.
Note 1. This time of day is the time of the original analog to digital conversion, and should not be changed thereafter. Midnight is represented by all 0’s. In order to convert this sample code into correct time, the sampling frequency must be known accurately.	
Byte 22	
	Flag bits used to indicate if the contents of the channel status data are reliable. If the specified bytes are reliable then the associated bits are set to 0. If the bytes are unreliable, the associated bits are set to 1.
Bits 0–3	Reserved. Set to 0.
Bit 4	Bytes 0 to 5
Bit 5	Bytes 6 to 13
Bit 6	Bytes 14 to 17
Bit 7	Bytes 18 to 21
Byte 23	
	Channel status data Cyclic Redundancy Check Character (CRCC).
Bits 0–7	The CRCC allows the receiver to check for correct reception of the bytes 0 through 22 of the channel status block. It is generated by $G(x) = x^8 + x^4 + x^3 + x^2 + 1$. If a “minimum” implementation is done, this will default to all 0’s. The AES3 Standard provides further information on how to calculate this.

Cabling to be used for AES3 is specified as 110 Ω balanced twisted pair with shield. The impedance must be held over a frequency range from 100 kHz to 128 times the maximum frame rate to be carried. The line driver and line receiver circuits must have an impedance of 110 Ω ±20% over the same frequency range. While the acceptable tolerance of the cable impedance is not specified, it is noted that tighter impedance tolerances for the cable, driver, and receiver will result in increased distance for reliable transmission, and for higher data rates. If a 32 kHz sampling rate mono signal were carried in single channel double sampling frequency mode, the interface frequency range would only extend to 2.048 MHz. If a 48 kHz sampling frequency two-channel signal were to be carried, the interface frequency range would extend to 6.144 MHz, or about the 6 MHz bandwidth commonly quoted for AES3. However, if a 192 kHz sampling frequency two-channel signal were to be carried, the interface frequency range would extend to 24.576 MHz. As you can see, some uses of AES3 can extend the frequency range far beyond 6 MHz. If you are using a mode that has extended interface frequency, make sure that the transmitter, interconnect system, and receiver are all designed to meet specifications over the entire frequency range in use.

When AES3 was originally introduced it was thought that ordinary analog audio shielded twisted pair cable would be acceptable for carrying AES3 digital audio, and indeed that is often the case for shorter distances. However, the impedance and balance of common audio cable vary widely, and it was quickly determined that purpose built AES3 cable performed significantly better for AES3 than ordinary analog audio cable. It was later determined that such AES3 rated cable often also performed significantly better as analog audio cable than ordinary cable, so today we commonly see AES3 rated cable used in both analog and digital applications.

While the AES3 Standard makes mention of interconnect lengths of “a few hundred meters,” in practice distances beyond about 100 m often require the use of equalization to compensate for losses in the cabling. If such equalization is used, it must never be applied to the transmitter, but only to the receiver.

As an alternative to purpose built AES3 rated digital audio cable, structured wiring meeting Category 5 or greater is acceptable. Such cabling can be of either shielded twisted pair (STP) or unshielded twisted pair (UTP) construction. To deliver satisfactory performance, only one cable type (Category 5 or higher STP, Category 5 or higher UTP, or AES3 digital audio rated)

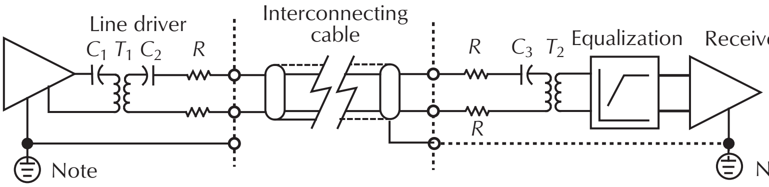


Figure 39-8. AES3 General Circuit Configuration. Note: If safety regulation so require, the signal grounds shown may also be tied to an electrical life safety ground.

may be used for the entire path from driver to receiver. If Category 5 or greater UTP is used, distances of 400 meters unequalized or 800 meters with equalization are possible.

39.3.4 Line Drivers

Just like AES3 cabling, line drivers are specified as having a balanced output with an impedance of 110 Ω ±20% over the entire frequency range from 100 kHz to 128 times the maximum frame rate. The driver must be capable of delivering an output level between 2–7 V (measured peak to peak) into a 110 Ω resistive termination directly across its output terminals with no cable present. The balance must be good enough so that any common mode output components are at least 30 dB lower in level than the balanced output signal.

The rise and fall times of the output, as measured between the 10–90% amplitude points, must be no faster than 5 ns, and no slower than 30 ns into a 110 Ω resistive termination directly across its output terminals with no cable present. A fast rise and fall time often improves the eye pattern at the receiver, but a slower rise and fall time often results in lower electromagnetic interference (EMI) radiated. Equipment must meet the EMI limits of the country in which it is used.

Equalization must not be applied at the driven end of the line.

39.3.5 Jitter

All digital equipment has the potential for introducing jitter, or small timing variations in the output signal. Extreme amounts of jitter can actually cause data errors. More moderate amounts of jitter may not change the actual data transmitted, but can lead to other ill effects. An ideal D/A would ignore the jitter on the incoming signal and perfectly produce the analog output based solely on the data carried. Unfortunately many real-world A/D and D/A converters are far from ideal, and allow jitter to change or modulate the output. Therefore keeping jitter low can have significant audible benefits.

AES3 divides the jitter at the output of a line driver into two parts intrinsic and pass through. The pass through portion of the jitter is due to jitter in the timing reference used. If such an external timing reference is used AES3 requires that there never be more than 2 dB of jitter gain at any frequency. The external timing reference may be derived from an AES3 input signal, or from a digital audio reference signal (DARS), which is an AES3 signal used as a clock reference as specified in AES11. If cascades of digital devices are built where each device uses as its clock reference the AES3 signal received from the previous device in the chain, it is possible for the pass through jitter to eventually increase the output jitter to an unacceptable level.

Many of today's better A/D and D/A converters provide jitter attenuation from the timing reference, Fig. 39-9.

Intrinsic jitter is measured with the equipment's own internal clock and with the equipment locked to an effectively jitter free external reference clock. Intrinsic jitter is measured through a minimum-phase one-pole high-pass filter whose –3 dB down point is 700 Hz, and which accurately provides that characteristic down to at least 70 Hz. The pass band of the filter has unity gain. Measuring at the transition zero crossings and through the filter the jitter must be less than 0.025 unit intervals (UI), Fig. 39-5.

39.3.6 Line Receivers

Just like AES3 cabling and line drivers, line receivers are specified as having a balanced output with an impedance of 110 Ω ±20% over the entire frequency range from 100 kHz to 128 times the maximum frame rate. The receiver must be capable of accepting an input level of 2–7 V (measured peak to peak). Early versions of AES3 required the receiver be able to accept 10 V. Only one receiver may be connected to an AES3 line. Early versions of AES3 permitted multiple receivers, but it became clear this was not good practice and the standard was modified.

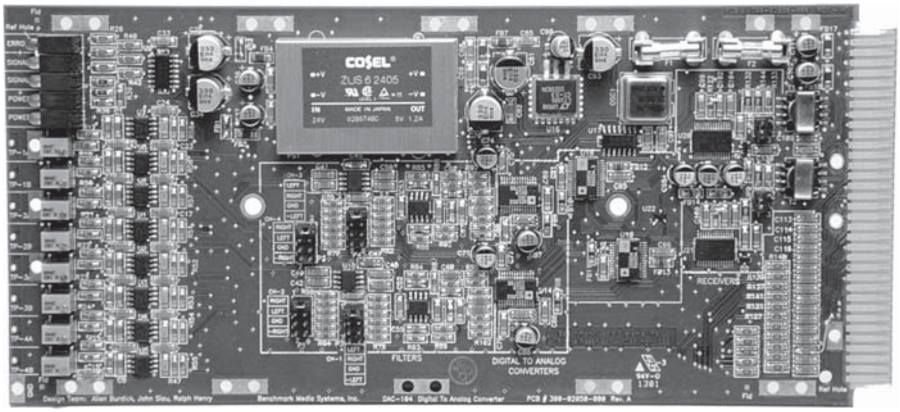


Figure 39-9. Benchmark Media Systems DAC-104 is a four channel 24 bit 96 kHz sampling rate D/A converter. An example of a high performance D/A, it provides jitter reduction of 100 dB at 1 kHz and 160 dB at 10 kHz. Total harmonic distortion plus noise (*THD* + *N*) is less than 0.00079% at –3 dB FS at any sampling rate and test frequency.

An AES3 receiver must correctly interpret data when a random data signal that is not less than $V_{min} = 200$ mV and $T_{min} = 0.5T_{nom}$, as shown in Fig. 39-10, is applied to the receiver.

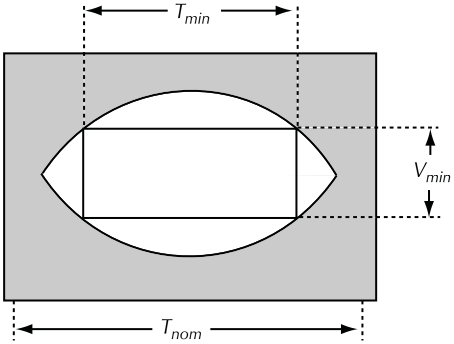


Figure 39-10. AES3 eye diagram. T_{nom} = 0.5 unit interval (UI) (see Fig. 39-5); $T_{min} = 0.5\ T_{nom}$; V_{min} = 200 mV. The eye diagram is one of the most powerful tools used to examine the quality of received data. The larger the open area of the eye the better. The limits shown are the most closed an eye should ever be for correct reception of the AES3 data.

If cable lengths of over 100 m are to be used, optional receiver equalization may be applied. The amount of equalization needed depends on the cable characteristics, length, and the frame rate of the AES3 signal. The AES3 Standard suggests that at a 48 kHz frame rate an equalizer with a boost that rises to a maximum of 12 dB at 10 MHz would be appropriate.

The receiver must introduce no errors due to the presence of common mode signals as large as 7 Vp at any frequency from dc to 20 kHz. This is not enough range to protect an AES3 receiver from the application of 10 Vdc digital phantom power as specified in the AES42 (AES3-MIC) Digital Interface for Microphones Standard.

The receiver must introduce no data errors from jitter that does not exceed 10 unit intervals (UI) at frequencies below 200 Hz decreasing to not exceeding 0.25 UI at frequencies over 8 kHz. Of course the recovered clock from such a high jitter signal may cause other problems, but at least the data must be decoded correctly.

39.3.7 AES3 Connectors

The connector for AES3 signals is what is commonly called the XLR, and is standardized in IEC 60268-12 as the circular latching 3 pin connector. Outputs are on male connectors and inputs are on female connectors just as in common analog usage of this same connector. The shield or ground connection is on pin 1, and the signal connections are on pins 2 and 3. With AES3 digital signals, the relative polarity of pins 2 and 3 is unimportant.

To avoid confusion with analog audio signal connectors, AES3 suggests that manufacturers label AES3 outputs “digital audio output,” or “DO;” and AES3 inputs “digital audio input,” or “DI.”

An alternative modified XLR connector has been proposed to help make clear that the signal on the connector is digital and not analog, and via a keying scheme reducing the chances of accidental interfacing of inputs and outputs that are incompatible. There has been much discussion in the AES about changing to this connector, either for all AES3 signals, or at least for AES42 (AES3-MIC) digital microphone signals, but no consensus has been reached. It should also be noted that since the analog audio bandwidth usually does not significantly exceed 20 kHz, and the AES3 spectrum does not go below 100 kHz, it is possible for a single cable to carry both an analog audio signal and an AES3 signal at the same time. The proposed modified XLR connector could also allow such a dual use condition.

If Category 5 or greater UTP or STP is used, the RJ45 connector must be used. Pins 4 and 5 of the RJ45 are the preferred pair, with pins 3 and 6 the suggested alternative pair. It is suggested that if adaptors from XLR to RJ45 are used, pin 2 of the XLR should connect to pin 5 (or other odd numbered pin) of the RJ45, and pin 3 of the XLR should connect to pin 4 (or other even numbered pin) of the RJ45 connector.

39.4 AES-3id

AES-3id is a variant on AES3 where the signal is carried over unbalanced 75 Ω coaxial cable instead of over 110 Ω balanced cable. It can allow the transmission of AES3 information over distances of up to 1000 m. Analog video distribution equipment and cable may often be suitable for transmission of AES-3id data. This of course is a great convenience in video facilities.

At distances of up to 300 m, receiver equalization may not be needed. Equalization must never be applied at the line driver end.

The AES-3id information document provides extensive tables and circuit diagrams showing active and passive circuits for AES-3id transmission. Canare, among others, sells passive adapters between 110 Ω balanced AES3 and 75 Ω unbalanced AES-3id.

AES-3id was written based on the assumption of the sampling rates specified in AES3-2003 and not on double or quadruple rates as are sometimes used today. The basic techniques of AES-3id should extend to these higher rates, however.

The following information is based on AES-3id-2001. It is always advisable to obtain the latest revision of the information document.

39.4.1 Line Driver

AES-3id line drivers must have an impedance of 75 Ω and exhibit a return loss in excess of 15 dB from 100 kHz to 6 MHz. Obviously if frame rates in excess of 48 kHz as allowed by AES3 were to be used, wider bandwidths would be required. Much but not all modern video gear will have the bandwidth to correctly handle higher sampling rates.

The peak to peak output voltage into a 75 Ω 1% tolerance resistor must be between 0.8 V and 1.2 V, with a dc offset not to exceed 50 mV. The rise and fall times should be between 30 ns and 44 ns. These output voltage, dc offset, and rise and fall times have been chosen for compatibility with analog video distribution equipment. Lower dc offset values are desirable for longer transmission distances.

39.4.2 Interconnect System

AES-3id cable must be 75 ±3 Ω over the range from 100 kHz to 6 MHz. It is to be equipped with BNC connectors as described in IEC 60169-8 but with an impedance of 75 Ω instead of 50 Ω.

39.4.3 Line Receiver

AES-3id line receivers must have an impedance of 75 Ω and exhibit a return loss in excess of 15 dB from 100 kHz to 6 MHz. The receiver must be capable of correctly decoding signals with input levels of 0.8 V and 1.2 V (measured peak to peak).

An AES-3id receiver must correctly interpret data when a random data signal that is not less than V_{min} =320 mV and $T_{min} = 0.5T_{nom}$ as shown in Fig. 39-11 is applied to the receiver. For reliable operation at distances in excess of 1000 m, a receiver that operates correctly with a V_{min} = 30 mV may be required.

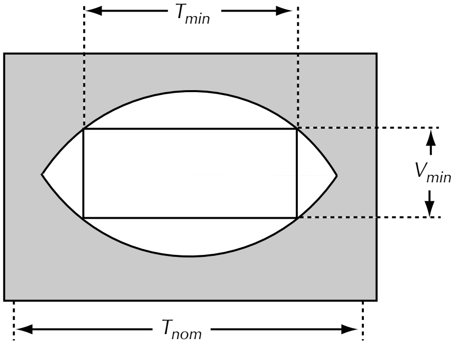


Figure 39-11. AES-3id Eye diagram. T_{nom} = 0.5 unit interval (UI) (see Fig. 39-5); T_{min} = 0.5 T_{nom}; V_{min} = 320 mV. The eye diagram is one of the most powerful tools used to examine the quality of received data. The larger the open area of the eye the better. The limits shown are the most closed an eye should ever be for correct reception of the AES-3id data.

39.5 AES42 (AES3-MIC)

AES42 (AES3-MIC) is a variant on AES3 designed to meet the needs of interfacing microphones that have direct digital outputs. The first most significant difference is that the transmitter and receiver use center tapped (on the cable side) transformer, which allow a digital phantom power (DPP) of +10 (+0.5, –0.1) Vdc at 250 mA to be supplied to the microphone. No more than 50 mVp-p ripple is allowed on the DPP. The microphone may draw no less than 50 mA, or more than 250 mA from the DPP, and may not present a load in excess of 120 nF to the DPP. The microphone must not be damaged by the application of any of the analog microphone phantom powers specified by IEC 61938 including common 48 V phantom. The techniques described by this standard may be applied to portable AES3 output devices other than microphones, however, AES42 only covers microphones.

Optionally a modulation from +10 to +12 V (resulting in a peak current of 300 mA) may be applied to the DPP for remote control purposes. This modulated signal thus travels in common mode from the AES3-MIC input back to the AES3-MIC microphone over the same cable that is carrying the AES3 audio data from the microphone to the AES3-MIC input. Because it is sent in common mode, the data rate must be far slower than that of AES3 to avoid interference. If the AES3 frame rate (FR) is 44.1 kHz or 48 kHz, the bit rate of the remote control signal is FR/64 bits per second (bit/s). For a FR of 88.2 kHz or 96 kHz the remote control bit rate is FR/128 bit/s. For a FR of 176.4 kHz or 192 kHz the remote control bit rate is FR/256 bit/s. As a result, the remote control bit rate is 750 bit/s if the AES3 FR is 48 kHz, 96 kHz, or 192 kHz, and 689.06 bit/s if the FR is 44.1 kHz, 88.2 kHz, or 176.4 kHz.

The remote control signals are sent as required, except if used for synchronization, in which case they will be sent on a regular basis of not less than six times per second.

The following information is based on AES42-2006. It is always advisable to obtain the latest revision of the standard.

39.5.1 Synchronization

There are two primary possible modes of operation for a microphone meeting the AES42 Standard, Fig. 39-12.

Mode 1 allows the microphone to free run at a rate determined by its own internal clock. No attempt is made to lock the microphone's clock rate to an external clock, and if such a lock is desired, sample rate conversion must be performed external to the microphone. This technique is the simplest way for an AES3-MIC microphone to operate, and does not require the use of the optional remote control signal.

Mode 2 uses the remote control signal to send data back to the microphone that allows its sampling rate to be varied, and phase locked to an external reference. The mode 2 microphone (or other AES3-MIC device) contains a voltage controlled crystal oscillator (VCXO), which has its frequency controlled by a digital to analog converter (DAC). The DAC receives control information via the remote control signal from the AES3-MIC receiving device. The receiving device compares the current sample rate of the microphone to the external reference and uses a phase locked loop (PLL) to generate a correction signal, which is sent back to the microphone. This results in the sampling rate of the microphone becoming frequency and phase matched to the reference signal. If multiple microphones or other AES3-MIC mode 2 sources are locked to the same reference, this has the additional advantage of providing a consistent and near zero phase relationship between the sampling times of the various sources. When multiple microphones sample correlated signals, for example, in stereo or multichannel recording techniques, this results in stable imaging.

If the receiver does not support mode 2 operation, the mode 2 microphone automatically reverts to mode 1 operation.

39.5.2 Microphone ID and Status Flags

AES42 defines the use of the user data channel in AES3 to optionally allow the microphone to identify itself and send back status information. Imagine the benefits in a complex setup of not having to worry which input a given mic is plugged into. The receiving device could use the microphone ID information to automatically route the microphone to the correct system input, no matter to which physical input it was connected.

39.5.3 Remote Control

AES42 defines three possible sets of remote control instructions, simple, extended, and manufacturer specific. If a device supports the extended instruction set, it must also support the simple instruction set. If a device supports at least the simple instruction set, it must have predetermined default settings it enters if no instructions are received on power up. If a device has switches on it, those will have priority over received instructions.

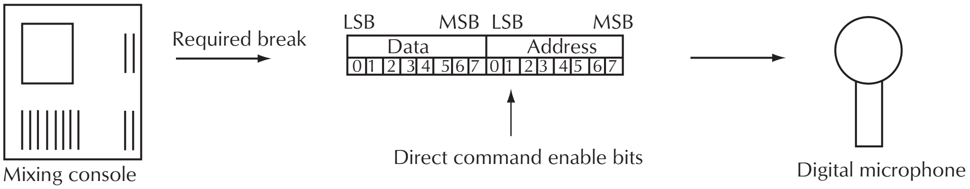


Figure 39-12. AES42 data format for the simple instruction set. Time flows from right to left in this diagram. Note that the MSB of each byte is sent first. After each 2 byte instruction there is a required break of 1 byte's worth of time. Addressing for the simple instruction set is contained in bits 0–2 of the address byte.

39.5.4 Simple Instruction Set

The simple instruction set is sent as a 2 byte signal with a minimum of 1 byte break between commands sent. Each byte is sent MSB first.

39.5.5 Direct Commands

Table 39-2 shows the direct commands.

Table 39-2. Direct Commands

Direct Command Address Byte		
Bits 0–2	Direct Command Enable bits	
Bit	0 1 2	
State	0 0 0	Identifies this command using the Extended Command set. See Extended Instruction Commands below.
	1 0 0	Direct Command 1 (low-cut filter, directivity control, preattenuation).
	0 1 0	Direct Command 2 (mute, limiter, signal gain).
	0 0 1	Direct Command 3 (synchronization).
	All other possible states of bits 0–2 are reserved and are not to be used unless defined by the AES in the future.	
Bits 3–7	Optional synchronization control word extension.	
Bit	3 4 5 6 7	
State	0 0 0 0 0	Default if synchronization control word extension not used.
	x x x x x	If the optional synchronization control word extension is used, bit 7 will be the MSB and bit 3 the LSB of the extension.
Direct Command 1 Data Byte		
Bits 0–1	Low-Cut Filter	
Bit	0 1	
State	0 0	No filter, default.
	1 0	Low-cut filter 1 (manufacturer defined).
	0 1	Low-cut filter 2 (manufacturer defined).
	1 1	Low-cut filter 3 (manufacturer defined).
Bits 2–5	Directivity	
Bit	2 3 4 5	
State	0 0 0 0	Manufacturer defined directivity, default.
	1 0 0 0	Omnidirectional.
	0 1 0 0	Increasing directivities
	0 0 1 0	through this state.
	1 0 1 0	Subcardioid.
	0 1 1 0	Increasing directivities
	1 1 1 0	Through this state.
	0 0 0 1	Cardioid.
	1 0 0 1	Increasing directivity.
	0 1 0 1	Supercardioid.
	1 1 0 1	Hypercardioid.
	0 0 1 1	Increasing directivities
	0 1 1 1	through this state.
	1 1 1 1	Figure of eight.
Bits 6–7	Preattenuation	
Bit	6 7	
State	0 0	No attenuation, default.
	1 0	Attenuation 1 (minimum).
	0 1	Attenuation 2.
	1 1	Attenuation 1 (maximum).
Direct Command 2 Data Byte		
Bit 0	Mute	
	0	Mute off, default.
	1	Mute on.
Bit 1	Limiter	
	0	Limiter disabled, default.
	1	Limiter enabled.
Bits 2–7	Gain	
Bit	2 3 4 5 6 7	
State	0 0 0 0 0 0	0 dB gain, default.
	0 0 0 0 0 1	+1 dB gain.
	x x x x x x	Increasing 1 dB per count.
	1 1 1 1 1 1	+63 dB gain.
Direct Command 3 Data Byte		
Bits 0–7	Synchronization	

Table 39-2. Direct Commands (Continued)

Bit	0 1 2 3 4 5 6 7	
State	0 0 0 0 0 0 0 0	Maximum negative tuning of VCXO.
	0 0 0 0 0 0 0 1	Center frequency of VCXO.
	1 1 1 1 1 1 1 1	Maximum positive tuning of VCXO.
Extended Instruction Commands		
Note that not all possible commands are defined.		
Extended Instruction Address Byte		
Bits 0–2	Direct Command Enable Bits.	
Bit	0 1 2	
State	0 0 0	Identifies this command uses the Extended Command set.
	1 0 0	Direct Command 1. See Direct Commands above.
	0 1 0	Direct Command 2. See Direct Commands above.
	0 0 1	Direct Command 3. See Direct Commands above.
	All other possible states of bits 0–2 are reserved and are not to be used unless defined by the AES in the future.	
Bits 3–7	Extended address bits.	
Bit	3 4 5 6 7	
State	0 0 0 0 0	Command 0, default.
	1 0 0 0 0	Command 4.
	0 1 0 0 0	Command 5.
	1 1 0 0 0	Command 6.
	0 0 1 0 0	Command 7.
	x x x x x	Commands follow in sequence.
	1 1 1 1 1	Command 34.
Extended Command 4 data Byte		
Bits 0–1	Light Control.	
Bit	0 1	
State	0 0	No light, default.
	1 0	Light 1.
	0 1	Light 2.
	1 1	Both lights.
Bits 2–3	Test Signal.	
Bit	2 3	
State	0 0	No test signal, default.
	1 0	Test signal 1 (under consideration).
	0 1	Test signal 2 (under consideration).
	1 1	Test signal 3 (under consideration).
Bit 4	ADC Calibration.	
	0	No calibration, default.
	1	Calibrate ADC
Bit 5	Reset	
	0	No reset, default.
	1	Reset.

Table 39-2. Direct Commands (Continued)

Bits 6–7	Microphone Status Data Page Request.	
Bit	6 7	
State	0 0	Page 0, default.
	1 0	Page 1.
	0 1	Page 2.
	1 1	Page 3.
Extended Command 5 Data Byte		
Bits 0–3	Dither and noise shaping.	
Bit	0 1 2 3	
State	0 0 0 0	No dither or noise shaping, default.
	x x x x	All other states are under consideration.
Bits 4–7	Sampling frequencies.	
Bit	4 5 6 7	
State	0 0 0 0	48 kHz, default.
	1 0 0 0	44.1 kHz.
	0 1 0 0	96 kHz multiple = 2.
	1 1 0 0	88.2 kHz multiple = 2.
	0 0 1 0	192 kHz multiple = 4.
	1 0 1 0	176.4 kHz multiple = 4.
	x x x x	All other states reserved.
Extended Command 6 Data Byte		
Bits 0–6	XY balance, (Notes 1, 2).	
Bit	0 1 2 3 4 5 6	
State	0 0 0 0 0 0 0	Left 0.5, Right 0.5, Center, default.
	1 1 1 1 1 1 0	Left 1.0, Right 0.0, Left (A) channel only.
	1 0 0 0 0 0 1	Left 0.0, Right 1.0, Right (B) channel only.
	0 0 0 0 0 0 1	Same as for 1 0 0 0 0 0 1.
Bits 0–v6	MS width (Notes 1, 2).	
Bit	0 1 2 3 4 5 6	
State	0 0 0 0 0 0 0	Mid 0.5, Side 0.5, Stereo, default.
	1 1 1 1 1 1 0	Mid 1.0, Side 0.0, Mono.
	1 0 0 0 0 0 1	Mid 0.0, Side 1.0, Difference only.
	0 0 0 0 0 0 1	Same as for 1 0 0 0 0 0 1.
Bit 7	XY or MS Select (Note 2).	
	0	XY stereo, default.
	1	MS stereo.

Note 1. Signed two's complement notation is used to encode the channel weights. Bit 6 is the sign extension = -2^0, bit 5 = 2^{-1} . . . Bit 0 = 2^{-6}.

Note 2. If XY is selected, then AES3 Channel 1 carries the Left audio channel, AES3 Channel 2 carries the Right audio channel, and bits 0–6 control the XY balance. If MS is selected, then AES3 Channel 1 carries the Mid or sum signal, AES3 Channel 2 carries the Side or difference signal, and bits 0–6 control the MS width.

Extended Command 7 Data Byte		
Bits 0–7	Equalization Curve Select.	
Bit	0 1 2 3 4 5 6 7	

Table 39-2. Direct Commands (Continued)

State	0 0 0 0 0 0 0 0	No equalization, default.
	x x x x x x x x	All other states, manufacturer specific equalization.
Extended Commands 8 through 32 Reserved.		
Extended Command 33 Manufacturer Specific Instruction begin.		
Extended Command 34 Manufacturer Specific Instruction end.		
Manufacturer Specific Instructions Are Under Consideration.		

39.5.6 Remote Control Pulse Structure

The remote control pulses are added to the DPP voltage and have a peak to peak amplitude of 2 ±0.2 V. They carry information in the form of pulse width modulation.

For AES3 frame rates (FR) of 48 kHz or multiples thereof the remote control data rate is 750 bit/s, while for FR of 44.1 kHz or multiples thereof the remote control data rate is 689 bit/s.

A logical 1 is represented by a pulse width of (7 × 64)/(8 FR), and must follow the preceding pulse at an interval of (1 × 64)/(8 FR). A logical 0 is represented by a pulse width of (1 × 64)/(8 FR), and must follow the preceding pulse at an interval of (7 × 64)/(8 FR). Thus in both cases the total time used by a bit is 64/FR, a byte is (8 × 64)/FR, and the combination of the command and data bytes is (16 × 64)/FR, if the FR is 44.1 kHz or 48 kHz.

It is possible that in the future an extended command byte may be added preceding the existing command and data bytes. In any case the entire sequence of extended command byte (if defined in the future), command byte, and data byte is sent with no interruptions in the flow of pulses.

The minimum interval between the end of one command and data bytes block and the beginning of the next is (8 × 64)/FR or a 1 byte interval. This allows detection of the end of the command and data bytes and for the data to be latched into the microphone.

The command byte is transmitted first, immediately followed by the data byte. Within each byte the MSB is transmitted first and the LSB last.

The rise and fall times of the pulses (measured from the 10% and 90% amplitude points) is to be 10 µs ±5 µs, over the entire specified load range of dc = 50–250 mA, C_{load} = 0–170 nF including the cable capacitance, Fig. 39-13.

39.5.7 Synchronization

A mode 2 AES3-MIC transmitter contains a VCXO and a DAC that set its operating frequency. The corresponding PLL resides in the AES3-MIC receiver. The receiver sends a regular stream of control voltage commands to the microphone using Direct Command 3 of the simple instruction set. The commands are repeated not less than once every $^1/_6$ s, and can have 8 to 13 bits of resolution. The ADC and DAC must have an accuracy of ±½ LSB, and be monotonic.

If on power up a mode 2 AES3-MIC transmitter does not see synchronization commands sent to it by the receiver, it will run in mode 1 at its default sampling rate or at the rate specified by the extended instruction set if supported. If while running a mode 2 AES3-MIC transmitter stops receiving synchronization commands, it should hold the last value of control voltage sent to it until synchronization commands are restored.

Mode 2 transmitters identify themselves to mode 2 receivers by means of a command that is part of the user data bits in the AES3 data stream. When a mode 2 capable receiver sees this signal it switches to mode 2 operation.

AES42 specifies the mode 2 AES3-MIC receiver characteristics for 48 kHz or 44.1 kHz operation. Since there is a linear relationship between comparison

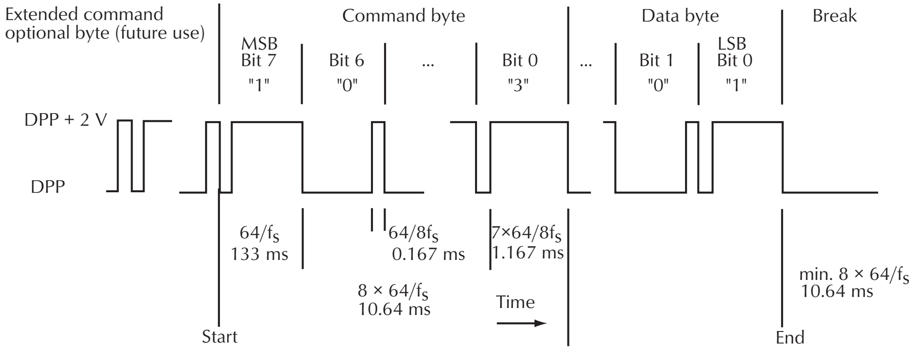


Figure 39-13. AES42 command and data byte bit structure at a 48 kHz frame rate (FR).

frequency and loop gain, operation at higher frequencies will require either frequency division down to 48 kHz or 44.1 kHz, or a corresponding reduction of the loop gain.

The phase comparator is a frequency-phase (zero-degree) type, and the PLL has a proportional, integrating, differentiating (PID) characteristic. The proportional constant *Kp* is 1 LSB at 163 ns time error (2.8° at 48 kHz). The integration time constant *Ki* is 1 LSB/s at 163 ns time error. The differential constant Kd is 1 LSB at 163 ns/s change of time error. The differential signal maximum gain is 8 LSB at 163 ns time error (fast change). The master reference clock must have an accuracy of ±50 ppm or better.

The AES3-MIC mode 2 transmitter must have a VCXO basic accuracy of ±50 ppm, a minimum tuning range of ±60 ppm + basic accuracy, a maximum tuning range of ±200 ppm, and a tuning slope that is positive with f_{max} for control data = 0xFF. The control voltage low pass filter has a dc gain of unity, a stage 1 filter that is first order with a corner frequency of 68 mHz (0.068 Hz) and maximum attenuation for frequencies greater than 10 Hz of 24 dB constant, and a stage 2 filter that is first order with a corner frequency of 12 Hz. Means may be used to raise the corner frequencies when the rate of tuning change is great. This will allow faster lockup on power on.

AES42 provides schematics showing how such a mode 2 control system might be implemented.

39.5.8 Microphone Identification and Status Flags

AES42 compliant transmitters may send status information to the receiver using the user data bit as defined in AES3. The channel status block start preamble is used to identify the start of blocks of 192 bits of user data. Each subframe contains user data. This allows different information to be sent in each subframe that is associated with that subframe.

In monophonic microphones where the audio data is repeated in both subframes, the user data must also be repeated.

Microphone status data is sent in MSB form in pages of 192 bits each. The pages are organized into 24 bytes. Byte 0 of all pages always contains the same data, including a page identifier, and time critical bits. This assures the delivery of the time critical bits no matter which page is being sent. Page 0 is sent continuously with each additional page being sent at least once per second. The receiver may request additional pages using the page request command in Extended Command data byte 4.

In order for the receiver to properly interpret the user data, the transmitter must set byte 1 bits 4–7 of the AES3 channel status data to 0 0 0 1. This indicates the user data bit is in use, and it is organized into 192 bit blocks starting with the AES3 subframe Z preamble.

39.5.8.1 Organization

All status bytes are sent MSB first, Table 39-3.

Table 39-3. Status Data Page

Status Data Page 0		
Status Byte 0—Starts All Status Data Pages		
Bits 0–2	Reserved.	
Bit	0 1 2	
State	0 0 0	Reserved. Must always be set to 0 0 0.
Bit 3	Mute	
	0	Not muted, default.
	1	Muted.
Bit 4	Overload.	
	0	No overload, default.
	1	Overload.
Bit 5	Limiter.	
	0	Limiter not active, default.
	1	Limiter active.
Bits 6–7	Page Identifier.	
Bit	6 7	
State	0 0	Status Page 0.
	1 0	Status Page 1.
	0 1	Status Page 2.
	1 1	Status Page 3, reserved.
Status Data Page 0 Byte 1—Microphone Configuration Echo		
Bits 0–1	Low-Cut Filter Status Echo.	
Bit	0 1	
State	0 0	No filter, default.
	1 0	Low-cut filter 1 (manufacturer defined).
	0 1	Low-cut filter 2 (manufacturer defined).
	1 1	Low-cut filter 3 (manufacturer defined).
Bits 2–5	Directivity Status Echo.	
Bit	2 3 4 5	
State	0 0 0 0	Manufacturer defined directivity, default.
	1 0 0 0	Omnidirectional.
	0 1 0 0	Increasing directivities
	0 0 1 0	through this state.
	1 0 1 0	Subcardioid.
	0 1 1 0	Increasing directivities
	1 1 1 0	through this state.
	0 0 0 1	Cardioid.

Table 39-3. Status Data Page (Continued)

	1 0 0 1	Increasing directivity.
	0 1 0 1	Supercardioid.
	1 1 0 1	Hypercardioid.
	0 0 1 1	Increasing directivities
	0 1 1 1	through this state.
	1 1 1 1	Figure of eight.
Bits 6–7	Preattenuation Status Echo.	
Bit	6 7	
State	0 0	No attenuation, default.
	1 0	Attenuation 1 (minimum).
	0 1	Attenuation 2.
	1 1	Attenuation 1 (maximum).
Status Data Page 0 Byte 2—Microphone Switch Monitoring		
Bits 0–4	Reserved	
Bit	0 1 2 3 4	
State	0 0 0 0 0	Reserved, must always be set to 0 0 0 0 0.
Bits 5–6	Call Buttons	
Bit	5 6	
State	0 0	No button pressed.
	1 0	Button 1 pressed.
	0 1	Button 2 pressed.
	1 1	Buttons 1 and 2 pressed.
Bit 7	Remote Off.	
	0	Remote parameter setting is enabled, default.
	1	Remote parameter setting is disabled.
Status Data Page 0 Byte 3—Microphone Remote Control Feature Indicator 1 (sound)		
Bit 0	EQ Curve Selection	
	0	Equalization curve not available, default.
	1	Equalization curve available, set using Extended Command Data Byte 7.
Bit 1	Balance-Width.	
	0	MS width or XY balance not available, default.
	1	MS width or XY balance available, set using Extended Command Data Byte 6, bits 0–6.
Bit 2	MS-XY Switch.	
	0	MS or XY selection not available, default.
	1	MS or XY selection available, set using Extended Command Data Byte 6, bit 7.
Bit 3	Limiter.	
	0	Limiter not available, default.
	1	Limiter available, set using Direct Command Data Byte 2, bit 1.
Bit 4	Gain Control.	
	0	Signal gain settings not available, default.
	1	Signal gain settings available, set using Direct Command Data Byte 2, bits 2–7.

Table 39-3. Status Data Page (Continued)

Bit 5	Low Cut Filter.	
	0	Low cut filter settings not available, default.
	1	Low cut filter settings available, set using Direct Command Data Byte 1, bits 0–1.
Bit 6	Pattern Control.	
	0	Directivity pattern settings not available, default.
	1	Directivity pattern settings available, set using Direct Command Data Byte 1, bits 2–5.
Bit 7	Attenuation.	
	0	Attenuation settings not available, default.
	1	Attenuation settings available, set using Direct Command Data Byte 1, bits 6–7.
Status Data Page 0 Byte 4—Microphone Remote Control Feature Indicator 2 (control)		
Bit 0	Mode 2 Synchronization.	
	0	External synchronization not available, default.
	1	External synchroByte 3.
Bit 1	Dither-Noise Shaping.	
	0	Dither and noise shaping not available, default.
	1	Dither and noise shaping available, set using Extended Command Data Byte 5, bits 0–3.
Bit 2	Multiple Sampling Frequency.	
	0	Multiple sampling frequencies not available, default.
	1	Multiple sampling frequencies available, set using Extended Command Data Byte 5, bits 4–7.
Bit 3	Light Control.	
	0	Light control selection not available, default.
	1	Light control selection available, set using Extended Command Data Byte 4, bits 0–1.
Bit 4	Test Signal.	
	0	Test signal selection not available, default.
	1	Test signal selection available, set using Extended Command Data Byte 4, bits 2–3.
Bit 5	ADC Calibrate.	
	0	ADC calibration function not available, default.
	1	ADC calibration function available, set using Extended Command Data Byte 4, bit 4.
Bit 6	Reset.	
	0	Reset function not available, default.
	1	Reset function available, set using Extended Command Data Byte 4, bit 5.
Bit 7	Mute.	
	0	Mute selection not available, default.
	1	Mute selection available, set using Direct Command Data Byte 2, bit 0.
Status Data Page 0 Byte 5—Reserved		
All bits of Page 0 Byte 5 are reserved and should be set to 0.		
Status Data Page 0 Byte 6—Wireless Microphone Status Flags		
Bits 0–4	Reserved.	

Table 39-3. Status Data Page (Continued)

Bit	0 1 2 3 4	
State	0 0 0 0 0	Reserved, must always be set to 0 0 0 0 0.
Bit 5	Squelch.	
	0	Receiver squelch inactive, default.
	1	Receiver squelch active.
Bit 6	Link Loss.	
	0	RF link is operating, default.
	1	RF link not operating.
Bit 7	Low Battery.	
	0	No low battery condition, default.
	1	Low battery condition.

Note 1. This byte is only used for wireless microphones, wired microphones should set all bits to 0.

Status Data Page 0 Byte 7—Wireless Microphone Battery Status

Bits 0–1	Reserved.	
Bit	0 1	
State	0 0	Reserved, must always be set to 0 0.
Bits 2–5	Battery Charge Proportion.	
Bit	2 3 4 5	
State	0 0 0 0	100%, default.
	1 0 0 0	90%.
	0 1 0 0	80%.
	1 1 0 0	70%.
	0 0 1 0	60%.
	1 0 1 0	50%.
	0 1 1 0	40%.
	1 1 1 0	30%.
	0 0 0 1	20%.
	1 0 0 1	10%.
	0 1 0 1	0%.
	x x x x	All other states reserved.
Bits 6–7	Battery Type.	
Bit	0 1	
State	0 0	Not indicated, default.
	1 0	Battery type is primary cell.
	0 1	Battery type is rechargeable.
	1 1	Reserved.

Note 1. Microphones supporting low battery indication use byte 6 bit 7. Microphones supporting battery charge indication must use both byte 7 bits 2 – 5, and byte 6 bit 7.

Status Data Page 0 Byte 8—Wireless Microphone Error Handling Flags

Bits 0–2	Reserved.	
Bit	0 1 2	
State	0 0 0	Reserved, must always be set to 0 0 0.

Table 39-3. Status Data Page (Continued)

Bits 3–4	Error Concealment	
Bit	3 4	
State	0 0	Error concealment not in use, default.
	1 0	Error concealment in use.
	0 1	Reserved.
	1 1	Reserved.
Bits 5–7	FEC Capacity.	
Bit	5 6 7	
State	0 0 0	FEC capacity used 0%, default.
	1 0 0	FEC capacity used 20%.
	0 1 0	FEC capacity used 40%.
	1 1 0	FEC capacity used 60%.
	0 0 1	FEC capacity used 80%.
	1 0 1	FEC capacity used 100%.
	0 1 1	FEC capacity overloaded.
	1 1 1	Reserved.

Status Data Page 0 Bytes 9–23—Reserved

All bits of Page 0 Bytes 9–23 are reserved and should be set to 0.

Status Data Page 1 Bytes 1–12—Manufacturer Identification

Manufacturer identification information should be sent in 7 bit ASCII form using only printable characters in the range of 00, and 20 to 7E Hex, and starting in byte 1. The use of nonprintable characters in the range of 01 to 1F Hex is not allowed. Each byte has a 0 of reserved usage in bit 7, followed by the MSB of the ASCII code in bit 6, through the LSB in bit 0. This allows 12 characters for manufacturer identification. Fill any unused bytes with zeros.

Status Data Page 1 Bytes 13–20—Microphone Model Identification

Microphone model identification information should be sent in 7 bit ASCII form using only printable characters in the range of 00, and 20 to 7E Hex, and starting in byte 13. The use of nonprintable characters in the range of 01 to 1F Hex is not allowed. Each byte has a 0 of reserved usage in bit 7, followed by the MSB of the ASCII code in bit 6, through the LSB in bit 0. This allows 8 characters for microphone model identification. Fill any unused bytes with zeros.

Status Data Page 1 Bytes 21–23—Reserved

All bits of Page 1 Bytes 21–23 are reserved and should be set to 0.

Status Data Page 2 Bytes 1–8—Microphone Serial Number

Microphone serial number information should be sent in 7-bit ASCII form using only printable characters in the range of 00, and 20 to 7E Hex, and starting in byte 1. The use of nonprintable characters in the range of 01 to 1F Hex is not allowed. Each byte has a 0 of reserved usage in bit 7, followed by the MSB of the ASCII code in bit 6, through the LSB in bit 0. This allows 12 characters for microphone serial numbers. Fill any unused bytes with zeros.

Status Data Page 2 Byte 9—Microphone Hardware Revision Main Counter

The information is sent as two binary coded decimal (BCD) digits. The L-nibble or lower nibble is sent with its LSB in bit 0, and MSB in bit 3. The U-nibble or upper nibble is sent with its LSB in bit 4, and MSB in bit 7. Numbers less than 10 are coded with a leading 0 in the U-nibble.

Status Data Page 2 Byte 10—Microphone Hardware Revision Index Counter

Table 39-3. Status Data Page (Continued)

The information is sent as two binary coded decimal (BCD) digits. The L-nibble or lower nibble is sent with its LSB in bit 0, and MSB in bit 3. The U-nibble or upper nibble is sent with its LSB in bit 4, and MSB in bit 7. Numbers less than 10 are coded with a leading 0 in the U-nibble.
Bytes 9 and 10 together represent the entire hardware revision number with a range of 00.00 to 99.99. Byte 9 is the integer portion while byte 10 is the fractional portion.

Status Data Page 2 Byte 11—Microphone Software Revision Main Counter

The information is sent as two binary coded decimal (BCD) digits. The L-nibble or lower nibble is sent with its LSB in bit 0, and MSB in bit 3. The U-nibble or upper nibble is sent with its LSB in bit 4, and MSB in bit 7. Numbers less than 10 are coded with a leading 0 in the U-nibble.

Status Data Page 2 Byte 12—Microphone Software Revision Index Counter

The information is sent as two binary coded decimal (BCD) digits. The L-nibble or lower nibble is sent with its LSB in bit 0, and MSB in bit 3. The U-nibble or upper nibble is sent with its LSB in bit 4, and MSB in bit 7. Numbers less than 10 are coded with a leading 0 in the U-nibble.
Bytes 11 and 12 together represent the entire software revision number with a range of 00.00 to 99.99. Byte 11 is the integer portion while byte 12 is the fractional portion.

Status Data Page 2 Bytes 13–23—Reserved

All bits of Page 2 Bytes 13–23 are reserved and should be set to 0.

39.5.9 XLD Connector for AES3-MIC Applications

AES42 describes but does not require the use of a new connector called the XLD for AES3-MIC applications. It is a variant on the common XLR-type connector with the addition of keying. This connector was designed by Neutrik and is not in production as of this writing.

The reason for this connector is to prevent accidental interconnection between analog and digital circuits, and in particular between analog and AES3-MIC circuits. The keying was designed to be field removable so if someone were required to use the same cabling for both analog and digital applications that would still be possible, Figs. 39-14–Fig. 39-18.

A number of those involved with the AES42 effort felt that some such connector that prevents accidental interconnections is mandatory given the use of the relatively high current DPP. They feared it might damage analog outputs to which it might be accidentally connected. Others favored a different connector than the XLR because of the years of experience troubleshooting mismatched analog and digital circuit connections. It was pointed out that this issue would be better addressed by the group responsible for AES3 revisions. There was also a group that felt that given the large amount of existing infrastructure using the XLR for AES3 circuits, any change was unacceptable. As a compromise, the standard was issued with the proposed new XLD connector described but not required.

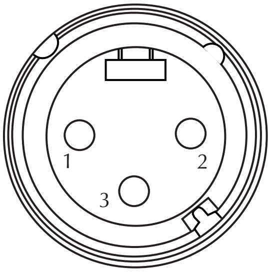


Figure 39-14. Fully coded male cable variety. This is a modified XLR type of connector with removable keying that could be used to prevent accidental interconnections between analog and AES3 fully coded female XLD.

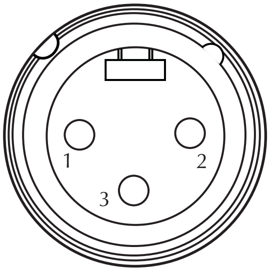


Figure 39-15. AES42 proposed XLD connector of the half-coded male cable variety. This is a modified XLR type of connector. It will mate with a standard female XLR or a fully coded female XLD.

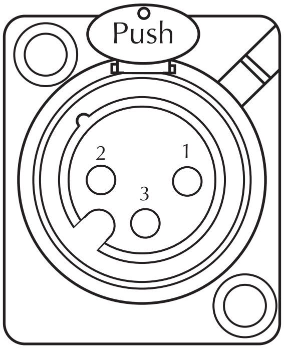


Figure 39-16. AES42 proposed XLD connector of the half-coded female chassis variety. This is a modified XLR type of connector. It will mate with a standard male XLR or a half- or fully coded male XLD.

According to AES42 the issue of what connector to use for AES3-MIC applications is under consideration by the AES Standards Committee. Meanwhile, the XLR-3 as currently specified for AES3 may be used. If a new connector is selected for AES3 applications, then

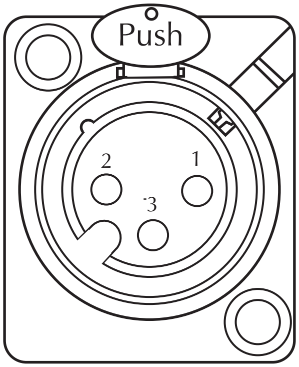


Figure 39-17. AES42 proposed XLD connector of the fully coded female chassis variety. This is a modified XLR type of connector with removable keying that could be used to prevent accidental interconnections between analog and AES3 or analog and AES3-MIC circuits. It will only mate with a fully coded male XLD.

that connector will be allowed for AES3-MIC as well. If a connector different from the XLR is desired currently, then that connector must be the XLD, Fig. 39-18.

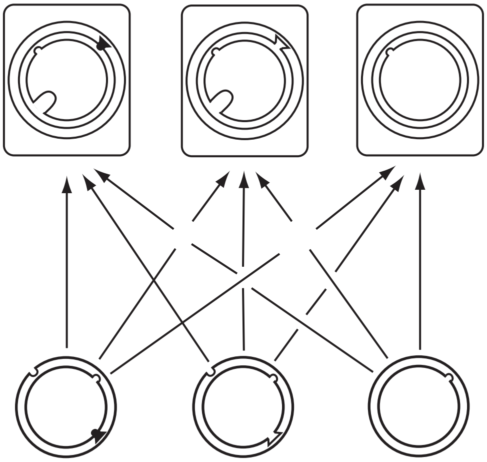


Figure 39-18. AES42 proposed XLD connection logic showing which varieties of XLD and XLR will mate. Note that if the fully coded varieties are used, it will not be possible to mate them with ordinary XLR connectors. If the keys are removed resulting in the half-coded variety, then the XLD will mate with either the XLR or the fully coded XLD. This was designed to satisfy those who insisted on maintaining compatibility with the XLR.

The black-white-black-white pattern on the zebra ring and associated wiring and the bumps on the surface of the ring serve to identify a connector that carries an AES3 digital audio signal, and that may be carrying DPP, Fig. 39-19. It also indicates that associated circuitry will not be damaged by the application of DPP. Cables so marked are designed to carry AES3 digital signals.

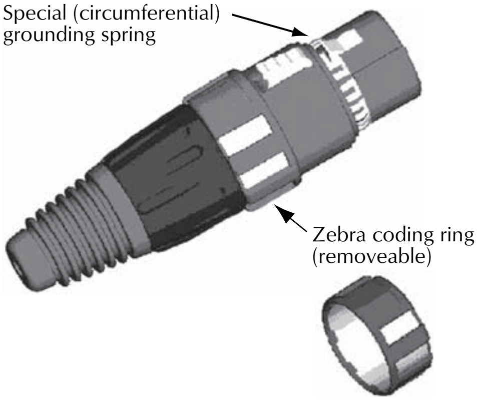

Special (circumferential) grounding spring
Zebra coding ring (removeable)

Figure 39-19. FAES42 proposed XLD zebra coding ring. The standard requires that if the XLD is used for AES3-MIC or AES3 signals the zebra ring also must be used. The zebra ring has black and white stripes and a bumpy surface to provide tactile feedback to the user. The grounding spring shown provides better grounding at RF frequencies and helps prevent RFI from entering or leaving the connector.

39.6 IEC 60958 Second Edition

This standard is based on three different sources, the AES3 and EBU Tech. 3250-E professional digital audio interconnection standards, and the consumer digital interface specification from Sony and Phillips (SPDIF).

The standard is broken into four parts, 60958-1 Ed, which contains general information on the digital interface; 60958-2 (unchanged from the first edition) on the serial copy management system: 60958-3 Ed2, which contains the consumer interface specific information: and 60958-4 Ed2, which contains information on the professional interface.

Since the professional interface is covered under the section on AES3 above, in this section we will only review how the consumer interface specified in 60958-3 differs from AES3.

Table 39-4 is based on Edition 3 of IEC 60958. It is always advisable to obtain the latest revision of the standard.

39.6.1 Electrical and Optical Interface

Two types of interface are specified, unbalanced electrical and optical fiber.

Three levels of timing accuracy are specified and indicated in the Channel Status. Level I is the high accuracy mode, requiring a tolerance of ±50 ppm. Level II is the normal accuracy mode, requiring a tolerance of ±1000 ppm. Level III is the variable pitch mode. An exact frequency range is under discussion, but may be ±12.5%.

By default, receivers should be capable of locking to signals of a Level II accuracy. If a receiver has a narrower locking range, it must be capable of locking to signals of a Level I accuracy, and must be specified as a Level I receiver. If a receiver is capable of normal operation over the Level III range, it should be specified as a Level III receiver.

Table 39-4. IEC 60958 Edition 2 Standard

Channel Status General Format		
Byte 0		
Bit 0	0	Contents of the channel status block conform to IEC 60958-3 “consumer use” Standard.
	1	Contents of the channel status block at to the AES3 “professional use” Standard. Ignore the rest of this table. (See Note 1.)
Bit 1	0	Audio words consist of linear PCM samples.
	1	Audio words consist of something other than linear PCM samples.
Bit 2	0	Software copyrighted. (See note 2.)
	1	Software copyright not claimed.
Bits 3–5	Additional format information, depending on the state of bit 1.	
If bit 1 = 0, linear PCM mode:		
Bit	3 4 5	
State	0 0 0	2 audio channels not using pre-emphasis.
	1 0 0	2 audio channels using 50/15 µs pre-emphasis.
	0 1 0	Reserved (for 2 audio channels using pre-emphasis).
	1 1 0	Reserved (for 2 audio channels using pre-emphasis).
	All other possible states of bits 3–5 are reserved and are not to be used unless defined by the IEC in the future.	
If bit 1 = 1, other than linear PCM mode:		
Bit	3 4 5	
State	0 0 0	Default state.
	All other possible states of bits 3–5 are reserved and are not to be used unless defined by the IEC in the future.	
Bits 6 – 7	Channel Status Mode.	
Bit	6 7	
State	0 0	Mode 0, Consumer use.
	All other possible states of bits 6–7 are reserved and are not to be used unless defined by the IEC in the future.	
Note 1. Other than the use of the Channel Status block of information, the rest of the data format is identical between the AES3 “professional use” Standard and the IEC 60958-3 “consumer use” Standard. The electrical format is different, however. For these reasons it should never be assumed that a “consumer use” receiver would function correctly with a “professional use” transmitter, or vice-versa.		
Note 2. If the copyright status is unknown for this application, the state of this bit may alternate at a rate between 4 Hz and 10 Hz.		
Channel Status Format for Consumer Use Digital Audio		
If Byte 0 bit 1, and bits 6–7 are all 0, then the following applies.		
Byte 1—Category Code		
Contains the category code indicating the type of equipment generating the signal. Category codes are given in the annexes to the Standard. Bit 0 contains the LSB and bit 7 the MSB. Used in conjunction with the copyright bit to control allowable copying of material.		
Byte 2—Source and Channel Number.		
Bits 0–3	Source Number	
Bit	0 1 2 3	
State	0 0 0 0	Don’t care.
	1 0 0 0	1.
	0 1 0 0	2.
	1 1 0 0	3.
	- - - -	
	1 1 1 1	15.
Bits 4–7	Audio Channel Number.	
Bit	4 5 6 7	
State	0 0 0 0	Don’t care.
	1 0 0 0	A (Left channel of stereo).
	0 1 0 0	B (Right channel of stereo).
	1 1 0 0	C.
	- - - -	
	1 1 1 1	O.
Byte 3—Sampling Frequency and Clock Accuracy.		
Bits 0–3	Sampling Frequency	
Bit	0 1 2 3	
State	0 0 0 0	44.1 kHz.
	0 1 0 0	48 kHz.
	1 1 0 0	32 kHz.
	All other possible states of bits 0–3 are reserved and are not to be used unless defined by the IEC in the future.	
Bits 4–5	Clock Accuracy.	

Table 39-4. IEC 60958 Edition 2 Standard (Continued)

Bit	4 5	
State	0 0	Level II.
	1 0	Level I.
	0 1	Level III.
	1 1	Reserved.
Bits 6–7	Reserved.	
Byte 4—Word Length		
Bit 0	Maximum audio word length.	
	0	Maximum 20 bit audio words.
	1	Maximum 24 bit audio words.
Bits 1–3	Encoded audio word length.	
Bit	1 2 3	Audio word length if bit 0 indicates max. 20 bit length.
State	0 0 0	Length not indicated, default.
	1 0 0	16 bits.
	0 1 0	18 bits.
	0 0 1	19 bits.
	1 0 1	20 bits.
	0 1 1	17 bits.
	All other possible states of bits 1–3 are reserved and are not to be used unless defined by the IEC in the future.	
Bits 4–7	Reserved.	

Note 1. If the auxiliary sample bits are not used they should be set to 0.

Note 2. Generally user data is not used and all bits are set to 0.

Note 3. Channel status is identical for all channels, with the exception of the channel number if not set to all zeros (don't care).

39.6.1.1 Unbalanced Line

Connecting cables are unbalanced, shielded, with an impedance of 75 Ω ±26.25 Ω over the frequency range from 100 kHz to 128 times the maximum frame rate.

The line driver has an impedance of 75 Ω ±15 Ω at the output terminals over the frequency range from 100 kHz to 128 times the maximum frame rate. The output level is 0.5 ±0.1 Vp-p, measured across a 75 ±1% Ω resistor across the output terminals without any cable connected. The rise and fall times measured between the 10% and 90% amplitude points should be less than 0.4 UI. The jitter gain from any reference input must be less than 3 dB at all frequencies.

The receiver should be basically resistive with an impedance of 75 Ω ±5% over the frequency range from 100 kHz to 128 times the maximum frame rate. It should correctly interpret the data of a signal ranging from 0.2 to 0.6 Vp-p.

The connector for inputs and outputs is described in 8.6 of Table IV of IEC 60268-11, and popularly known as the RCA connector. Male plugs are used at both ends of the cable. Manufacturers should clearly mark digital inputs and outputs.

39.6.1.2 Optical Connection

This is specified in IEC 61607-1 and IEC 61607-2, and popularly known as the TOSLINK connector.

39.7 AES10 (MADI)

The AES10 Standard describes a serial multichannel audio digital interface, or MADI. The abstract says it uses an asynchronous transmission scheme, but the overall protocol is better described as isochronous. It is based on the AES3 Standard, but allows thirty two, fifty six, or sixty four channels of digital audio at a common sampling rate in the range of 32 to 96 kHz, with a resolution of up to 24 bits to be sent over a single 75 Ω coaxial cable at distances up to 50 m. Transmission over fiber is also possible. Like the other schemes we have looked at it only allows one transmitter and one receiver.

Table 39-5 is based on AES10-2003. It is always advisable to obtain the latest revision of the standard.

MADI used the bit, block, and subframe structure of AES3 with the exception of the subframe preambles. Instead it substitutes four bits according to Table 39-5.

MADI sends all its active channels in consecutive order starting with channel zero. Each active channel has the active channel bit set to 1. Inactive channels must have all their bits set to 0 including the channel active bit. Inactive channels must always have higher channel numbers than any active channel.

The channels are transmitted serially using a nonreturn-to-zero inverted (NRZI) polarity free coding. Each 4 bits of the data are turned into 5 bits before encoding.

Table 39-5. AES10 MADI

Bit	Name	Description	Sense
0	MADI channel 0	Frame synchronization bit	1 = true
1	MADI channel active	Channel active bit	1 = true
2	MADI channel A or B	AES3 sub-frame 1 or 2	1 = sub-frame 2

Table 39-5. AES10 MADI (Continued)

3	MADI channel block sync	Channel block start	1 = true
4–27	AES3 audio data bits	(bit 27 is MSB)	
28	AES3 V	Validity bit	
29	AES3 U	User data bit	
30	AES3 C	Status data bit	
31	AES3 P	Parity bit (excludes bits 0–3)	Even

Each 32 bit channel data is broken down into eight words of 4 bits each following this scheme:

Word	Channel Data Bits			
0	0	1	2	3
1	4	5	6	7
2	8	9	10	11
3	12	13	14	15
4	16	17	18	19
5	20	21	22	23
6	24	25	26	27
7	28	29	30	31

The 4 bit words are turned into 5 bit words as follows:

4 Bit Data	5 Bit Encoded Data
0000	11110
0001	01001
0010	10100
0011	10101
0100	01010
0101	01011
0110	01110
0111	01111
1000	10010
1001	10011
1010	10110
1011	10111
1100	11010
1101	11011
1110	11100
1111	11101

The now 5 bit words are transmitted (left to right) as follows:

Word	Channel Link Bits				
0	0	1	2	3	4
1	5	6	7	8	9
2	10	11	12	13	14
3	15	16	17	18	19
4	20	21	22	23	24
5	25	26	27	28	29
6	30	31	32	33	34
7	35	36	37	38	39

Unlike the coding used for AES3, this coding allows dc on the link.

AES10 uses a synchronization symbol, 11000 10001 transmitted left to right, which is inserted at least once per frame to ensure synchronization of the receiver and transmitter. There are no defined locations for the insertion of this symbol, but it may only be inserted at the 40 bit boundaries between data words. Enough synchronization symbols should be interleaved between channels transmitted, and after the last channel has been transmitted, to fill up the total link capacity.

39.7.1 NRZI Encoding

The 5 bit link channel data is encoded using NRZI polarity free encoding. Each high bit is converted into a transition from the bit before, while each low bit results in no transition. In other words a 1 turns into a 1 to 0 or 0 to 1 transition, while a 0 results in a static 1 or 0.

39.7.2 Sample Frequencies and Rates

MADI allows operation with sampling rates in any of three ranges:

- 32 to 48 kHz, ±12.5%, 56 channels.
- 32 to 48 kHz nominal, 64 channels.
- 64 to 96 kHz, ±12.5%, 28 channels from the nominal frequency.

Higher sampling rates such as 192 kHz require the use of multiple audio channels per sample.

Data is transmitted across the link at a constant 125 megabits per second irrespective of the number of channels in use. The data transfer rate is 100 megabits per second. The difference is due to the 4 data bit to 5 link bit encoding used.

Actual data rates used will vary. fifty six channels at 48 kHz + 12.5% or 28 channels at 96 kHz + 12.5% results in a data rate of 96.768 megabits per second, while 56 channels at 32 kHz – 12.5% results in a data rate of 50.176 megabits per second.

39.7.3 Synchronization

Unlike AES3, MADI does not carry synchronization information. Therefore a separate AES3 signal must be provided to both the transmitter and receiver for synchronization purposes.

A MADI transmitter must start each frame within 5% of the sample period timing of the external reference signal. A MADI receiver must accept frames that start within 25% of the sample period timing of the external reference signal.

39.7.4 Electrical Characteristics

Either 75 Ω coax or optical fiber is allowed for the transmission media. Optical interfacing is described below.

The line driver has an impedance of 75 Ω ±2 Ω average output level when terminated into 75 Ω is 0 V ±0.1 V. The peak-to-peak output voltage is between 0.3 V and 0.6 V into 75 Ω. Rise and fall times between the 20% and 80% amplitude points must be no longer than 3 ns, and no shorter than 1 ns with a relative timing difference to the average of the amplitude points of no more than ±0.5 ns.

Interestingly there is no input impedance specified for the receiver, although the example schematic shows a 75 Ω termination.

When a signal meeting the limits shown in Fig. 39-20 is applied to the input of a MADI receiver, it must correctly interpret it.

Cabling to interconnecting MADI devices must be 75 Ω ±2 Ω and have a loss of less than 0.1 dB/m over the range from 1–100 MHz. Cables are equipped with 75 Ω BNC-type male connectors and have a 50 m maximum length. Chassis connectors are female.

At the receive end of the cable the eye pattern must be no worse than what is shown in Fig. 39-20. Equalization is not allowed.

The cable shield must be grounded to the chassis at the transmitter. If the shield is not grounded directly to the chassis at the receiver, it must be grounded above 30 MHz. This can be achieved by capacitively coupling the shield to the chassis through a suitable low inductance capacitor of around 1.0 nF.

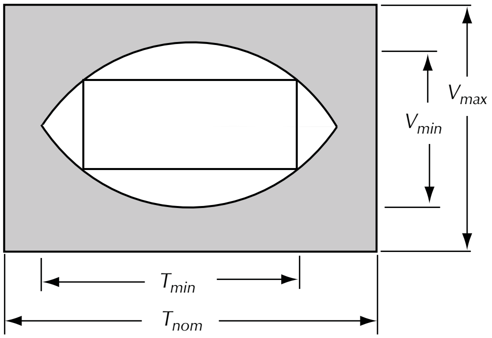


Figure 39-20. AES10 eye pattern for minimum and maximum input signals where t_{nom} = 8 ns; t_{min} = 6 s; V_{max} = 0.6 V; V_{min} = 0.15 V. The MADI receiver must correctly interpret signals within this eye diagram as applied to its input.

39.7.5 Optical Interface

Graded-index fiber with a core diameter of 62.5 nm, nominal cladding diameter of 125 nm, and a numerical aperture of 0.275 is to be used with ST1 connectors. This will allow links of up to 2000 m.

39.8 Soundweb

All of the interconnect schemes we have looked at so far have been point to point, and not networked. Soundweb is a good example of a simple yet useful networking scheme that is part of a family of signal-processing devices from BSS Audio. In the following discussion we will only consider the digital audio networking aspect of Soundweb. The chapter on virtual systems describes the sound-processing aspects of this family of products.

Unlike the standards-based digital audio interconnect methods discussed so far, the protocol for Soundweb is not published and is available only in products from a single manufacturer, BSS Audio. Nontheless it is in wide use, and needs to be examined from an applications viewpoint.

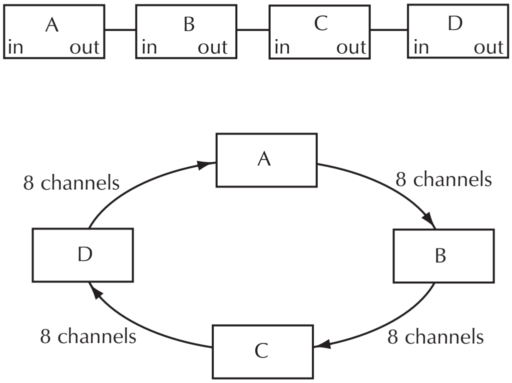
Each Soundweb component has network in and out connectors for interconnecting the devices. Most devices have just 1 in and 1 out, but their active hub has three in and three out connectors. Each of the six network connectors on a hub may be used to terminate one end of a chain. Virtual wiring inside the hub is then used to interconnect the 6 networks as desired.

An output is connected to an input with a Category 5 (Cat 5) data cable of up to 300 m in length. By using special fiber converters that distance can be extended to

2 km. The maximum number of units is dependent on the distance, but they say that usually up to fifteen units can be cascaded in a single chain.

Even though Soundweb uses Cat 5 cabling and the same "RJ-45" connectors as used for Ethernet networking, it is important to note that Soundweb is not using the Ethernet protocol, just the same cable and connectors as used by Ethernet.

Referring to Fig. 39-21, there are several things to note. First, the physical wiring must always be from output to input. Although signal flows in both directions over each cable, the output and input terminology shows the direction of primary signal flow. Second, even though the physical topology is a chain, the virtual signal flow topology is a loop. Up to eight audio channels may be passed from device to device over each virtual signal flow link. This requires planning since if, for example, you wanted to get a signal from B to A above, you would have to pass the signal from B to C, connect it inside C to an output channel, pass it from C to D, connect it inside D to an output channel, and then pass it from D back to A. As a result you are using one of the eight available channels on each of three links in order to pass the signal back one box. Planning your circuit topology carefully will reduce the need to send signals backwards.



Audio signal flow that results from physical interconnection shown

Figure 39-21. Soundweb. The physical interconnection is a chain, but the audio signal path is a loop.

39.9 Nexus

The Stage Tec Nexus is another proprietary digital audio networking system. It is a very high-quality system using fiber optic interconnections, providing a very flexible system. Redundant interconnections allow very high reliability. A very wide variety of input and output devices are available to insert into the Nexus frames. Both analog and digital inputs and outputs are available.

One of the most interesting aspects of the Nexus system is the ability of the programmable devices in a system to learn what they should be doing from other devices in the system. If a device fails in a running system, when the replacement device is plugged in, it determines what it should be doing from the other devices in the network. The instructions for all the devices are stored in multiple places in the system to enable this capability.

39.10 IEEE 1394 (FireWire)

FireWire is an attractive networking scheme since it provides both isochronous and asynchronous capabilities in the same network. However, this same diversity of capabilities that makes it attractive as a home networking scheme has also lead to a wide variety of incompatible protocols that are all carried over IEEE 1394.

Equipment from different manufacturers, and sometimes even different models from the same manufacturer are not compatible. James Snider, chairman of the 1394 Trade Association wrote in March 2001 that users of 1394 enabled equipment "want to know exactly what will work together and what will not, so they do not have unreasonable performance expectations." The 1394 Trade Association is working on this issue, however, the issues are not totally resolved.

FireWire has a maximum distance limitation for a given link of 4.5 m and has a bandwidth limitation of 400 MBit/s for the entire network. In this sense it is similar to an Ethernet repeater-based network. FireWire's proponents claim it can be cheaper than Ethernet, but market forces have dropped the cost of Ethernet and increased the capabilities to the point one could question if FireWire will ever catch up.

Lastly, 1394 has been mostly applied to consumer video and gaming application to this point. While it can certainly carry audio, there are few if any professional digital audio devices currently using this protocol.

39.11 Ethernet

Ethernet is the most common digital networking standard in the world, with over 50 million nodes installed. The huge volumes in the computer industry are constantly driving the price down and the capabilities upward. Many professional audio devices use Ethernet for control and programming, and both AMX and

Crestron have embraced it for audio/video system control.

Yet with all these advantages, Ethernet per se is poorly suited to carry real-time audio since it is by nature an asynchronous system. Kevin Gross of Peak Audio decided that there had to be a way to overcome the limitations of Ethernet for transmission of real-time information such as audio and video. His solution, called CobraNet®, was granted a patent, and has been licensed by many major professional audio companies for inclusion in their products including: Biamp, Creative Audio, Crest Audio, Crown, Digigram, EAW, LCS, Peavey, QSC, Rane, and Whirlwind.

More recent entrants into digital audio networking include Aviom, EtherSound, and Dante. All of the above make use of at least some portion of Ethernet technology.

Before we can examine these audio networking technologies, we need to get a good understanding of Ethernet.

39.11.1 Ethernet History

In 1972 Robert Metcalf and his colleagues at the Xerox Palo Alto Research Center (PARC) developed a networking system called the Alto Aloha Network to interconnect Xerox Altos computers. Metcalf changed the name to Ethernet in 1973. While the Altos is long gone, Ethernet has gone on to become the most popular networking system in the world, Fig. 39-22.

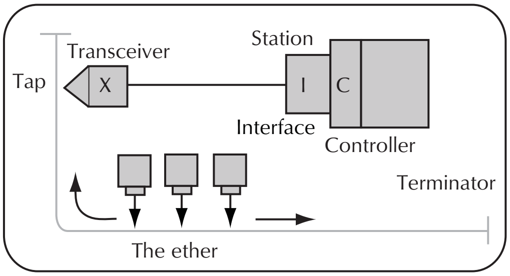


Figure 39-22. Robert Metcalf's first drawing of what became known as Ethernet.

This system had a number of key attributes. It used a shared media, in this case a common coaxial cable. This meant that the available bandwidth was shared among all the stations on the cable. If any one station transmitted, all the other stations would receive the signal. Only one station could transmit at any instant in time and get the data through uncorrupted. If more than one station attempted to transmit at the same time, this was called a collision, and resulted in garbled data.

In a shared media Ethernet network there will be collisions as a normal part of operation. As a result a mechanism for preventing collisions, detecting those it could not prevent, and recovering from them was required.

This mechanism is called *carrier-sense multiple access with collision detection* (CSMA/CD). In other words, while any station can transmit at any time (multiple access), before any station can transmit it has to make sure no other station is transmitting (carrier-sense). If no other station is transmitting, it can start to transmit. However, since it is possible for two or more stations to attempt to transmit at the same time, each transmitting station must listen for another attempted transmission at the same time (a collision). If a station transmitting detects a collision, the station transmits a bit sequence called a jam to insure all transmitting stations detect that a collision has occurred, then is silent for a random time before attempting to transmit again. Of course no retransmission can be attempted if another station is transmitting. If a second collision is detected, the delay before retransmission is attempted again increases. After a number of tries the attempt to transmit fails.

Now since the signals travel at the speed of light (approximately) down the coax cable, and since a station at one end of the cable has to be able to detect a collision with a transmission from a station at the other end of the cable, two requirements had to be imposed. First, there was a limitation on how long the cable could be. This was imposed to limit the time it could take for a transmission from a station at one end of the cable to reach the most distant station. Second, there was a minimum length imposed on the data packet transmitted. This made sure that stations that were distant from each other would have time to realize that a collision had occurred. If the cable were too long, or the packet were too short, and the stations at the ends of the cable were to both transmit at the same time, it would be possible for them to both finish transmitting before they saw the packet from the other station, and never realize that a collision had occurred, Fig. 39-23.

39.11.2 Ethernet Packet Format

Every Ethernet device in the world has a globally unique media access control or MAC address. Manufacturers apply to the Institute of Electrical and Elec-

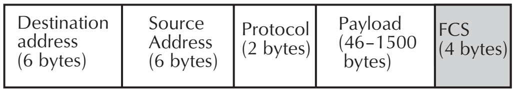


Figure 39-23. Ethernet packet format.

tronics Engineers (IEEE) and are assigned a block of MAC addresses for their use. Each manufacturer is responsible to make sure that each and every device it ships has a unique MAC address within that range. When it has used up 90% of its addresses it can apply for an additional block of addresses.

A MAC address is 48 bits, or 6 bytes long, which allows for 281,474,976,710,656 unique addresses. While Ethernet is extremely popular, we have not begun to run out of possible MAC addresses.

An Ethernet packet starts with the MAC address of the destination. That is followed by the MAC address of the station sending the packet. Next come 2 bytes called the EtherType number or protocol identifier, which identify the protocol used for the payload. Again the IEEE assigns these numbers. The protocol identifier assigned for CobraNet®, for example, is 8819 in hexadecimal notation.

The data payload can range from a minimum size of 46 bytes to a maximum size of 1500 bytes. The protocol identifier specifies the content of the payload and how it is to be interpreted. Data of less than 46 bytes must be extended or padded to 64 bytes, while data of more than 1500 bytes must be broken into multiple packets for transmission.

The frame check sequence (FCS) is a 4 byte long cyclic redundancy check (CRC) calculated by the transmitting station based on the contents of the rest of the Ethernet packet (destination address, source address, protocol, and data fields). The receiving station also calculates the FCS and compares it with the received FCS. If they match, the data received is assumed to have been received without corruption. There is a 99.9% probability that even a 1 bit error will be detected.

As you can see, the smallest possible Ethernet packet is 64 bytes long, and the longest is 1518 bytes long.

39.11.3 Network Diameter

The maximum allowable network diameter, Fig. 39-24, that will permit Ethernet's collision detection scheme to work is dependent on:

- The minimum packet size (64 bytes),
- The data rate (these last two together determine the time duration of the minimum size packet), and
- The quality of the cable (which determines the speed of propagation down the cable).

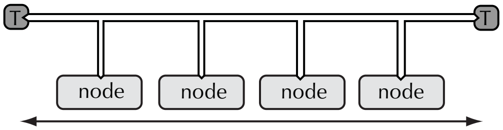


Figure 39-24. Ethernet maximum allowable network diameter.

39.11.4 Ethernet Varieties

In 1980 the IEEE standardized Ethernet as IEEE 802.3. This initial standard was based on the use of 10 mm 50 Ω coaxial cable. Many variations quickly appeared.

- 10Base5—This was the original Ethernet, also called thicknet or thick Ethernet because of the large diameter coaxial cable used. It ran at 10 MBit/s, baseband, with a maximum segment size of 500 m (hence 10Base5).
- 10Base2—Designed as a less expensive Ethernet, it was called thinnet or thin Ethernet due to the thinner RG-58 50 Ω coaxial cable used. It ran at 10 MBit/s, baseband, with a maximum segment length of 200 m.
- 1Base2—A slower variant of thinnet. It ran at 1 MBit/s, baseband, with a maximum segment length of 200 m.
- 10Broad36—Very rare, this ran over a RF cable plant similar to cable TV distribution systems, and was built with cable TV distribution components.

All of these variants suffer from a common problem. Since they use a shared media that had to physically connect to every station in the network, a problem at any point along the backbone could disable the entire network. Clearly a different approach was needed to protect the shared media from disruption.

In 1990 a new technology called 10Base-T was introduced to solve these problems. Instead of the vulnerable shared media being strung all over the entire facility, it was concentrated into a box called a repeater hub, Fig. 39-25. The media was still shared, but protected. It had an allowable network diameter of 2000 m, and used the same packet structure. Each station was connected to the hub with twisted pair cable, with two pairs used. One pair carried the signal from the station to the hub, while the other pair carried the signal from the hub to the station. Category 3 (Cat3) unshielded twisted pair (UTP) was used with transformer isolation at both ends of each pair. The maximum length of a single cable run was restricted to

100 m. Longer runs up to 2000 m were possible using a fiber version called 10BaseF. Cat3 cable was more durable and less expensive than the coax formerly required. Of greatest importance, a problem with a cable only affected a single station, and would not bring down the entire network.

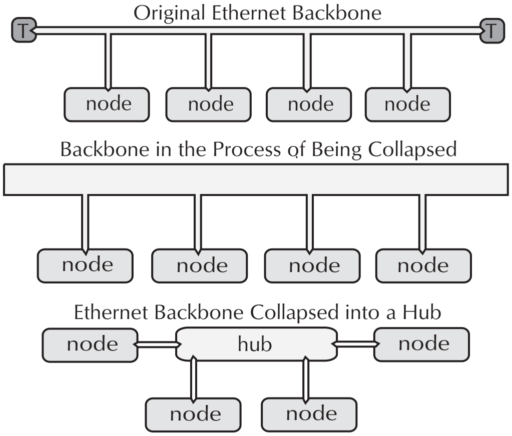

Original Ethernet Backbone
T
T
node
node
node
node
Backbone in the Process of Being Collapsed
node
node
node
node
Ethernet Backbone Collapsed into a Hub
node
hub
node
node
node

Figure 39-25. Transforming the Ethernet backbone into a repeater hub.

Since then, 100BaseT or Fast Ethernet has been introduced. It runs at ten times the data rate of 10Base-T, or 100 MBit/s. Since the data rate is ten times as high as 10Base-T, and the minimum packet size is the same, the maximum network diameter had to be reduced to 200 m. This is the most common form of Ethernet today although Gigabit Ethernet is catching up quickly. CobraNet® uses Fast Ethernet ports but can be transported over Gigabit Ethernet between switches.

Within Fast Ethernet there are several varieties. 100Base-T4 uses all 4 pairs of a Cat3 UTP cable. 100Base-TX uses 2 pairs of a Cat5 cable. This is the most common variety of Fast Ethernet. Both of these varieties allow single cable runs of 100 m. 100Base-FX uses multimode fiber, and allows single runs of up to 2000 m. A version of Fast Ethernet to run over single-mode fiber has not been Standardized, but many manufacturers sell their own versions, which allow distances of as much as 100,000 m in a single run.

Many Fast Ethernet devices sold today not only support 100Base-TX, but also 10Base-T. Such a dual speed port is commonly called a 10/100 Ethernet port. It will negotiate automatically with any Ethernet device hooked to it and connect at the highest speed both ends of the link support. The technique for this negotiation is described below.

Gigabit Ethernet is now available, and the price has dropped so much it is more and more replacing Fast Ethernet. As you might have guessed it runs at a rate ten times as fast as 100BaseT, or 1000 MBit/s. The first versions ran over optical fiber, but now a version that runs over Cat5 UTP cabling is available. It does, however, use all four pairs in the cable. Gigabit Ethernet increases the minimum packet size from 64 bytes to 512 bytes in order to allow the network diameter to stay at 200 m. Ethernet ports that support 10/100/1000 MBit/s speeds and auto-negotiate to match the highest speed the connected device can support are now common.

Within Gigabit Ethernet there are also several varieties. 1000Base-LX (L for long wavelength) can be used with either multimode or single-mode optical fiber. 1000Base-SX (S for short wavelength) is used with multimode fiber only. 1000Base-SX is less expensive than 1000Base-LX. 1000Base-LH (LH for long haul) is not an IEEE standard, but is supported by many manufacturers. Manufacturers make different versions depending on the distance to be covered. 1000Base-T runs over Cat5 cable using all four pairs. The maximum single cable run is 100 m.

A version of Ethernet that will run at ten times the speed of Gigabit Ethernet is available and the price has been dropping.

Several manufacturers power their products over Ethernet cabling. There is now an IEEE Standard for Power over Ethernet (PoE) and most manufacturers sending power to their products over the Ethernet cabling have gone to this standard.

Wireless Ethernet to the IEEE 802.11 Standard has become vary popular and inexpensive. It provides a variable data rate based on distance and environmental conditions. The best case data rate for 802.11n (the latest as of this writing) is 300 MBit/s, but typical data rates are closer to 74 MBit/s.

39.11.5 Ethernet Topology

The Ethernet topology shown in Fig. 39-25 as a collapsed backbone is commonly called a star topology, since every station connects back to the common hub. It is also permissible to tie multiple stars together in a star of stars, Figs. 39-26 and 39-27.

Using fiber to interconnect the stars can increase the distance between clusters of stars, Fig. 39-28.

39.11.6 Ethernet Equipment

This will become clearer when we examine the internal functions of the repeater hubs we have been talking

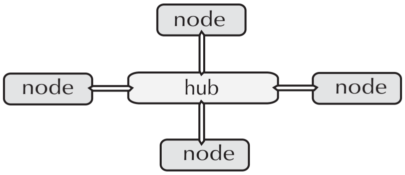


Figure 39-26. Ethernet star topology.

about. A hub has ports that connect either to stations or other hubs. Any data that comes in a port is immediately sent out all the other ports *except* the port it came in on, Fig. 39-29. An audio analogy would be a mix-minus system.

One of the factors keeping the size of the network from growing is that all of these star and star of stars topologies still have the same network diameter limitation. One way to build a bigger network is to isolate data in one star from that in another, and only pass between stars packets that need to reach stations in the other star. Collisions that occur in a given star are not passed to the other stars since only complete packets addressed to a station in the other star are passed on. This isolates each star into a collision domain of its own, so the network diameter limitation only applies within a given collision domain.

The device that provides this function between a pair of collision domains is called a bridge. As the technology became cheaper multiport bridges started to appear that were called switches. As switches become popular and bridges fade from use, you will sometimes see a bridge referred to as a two-port switch, Fig.39-30.

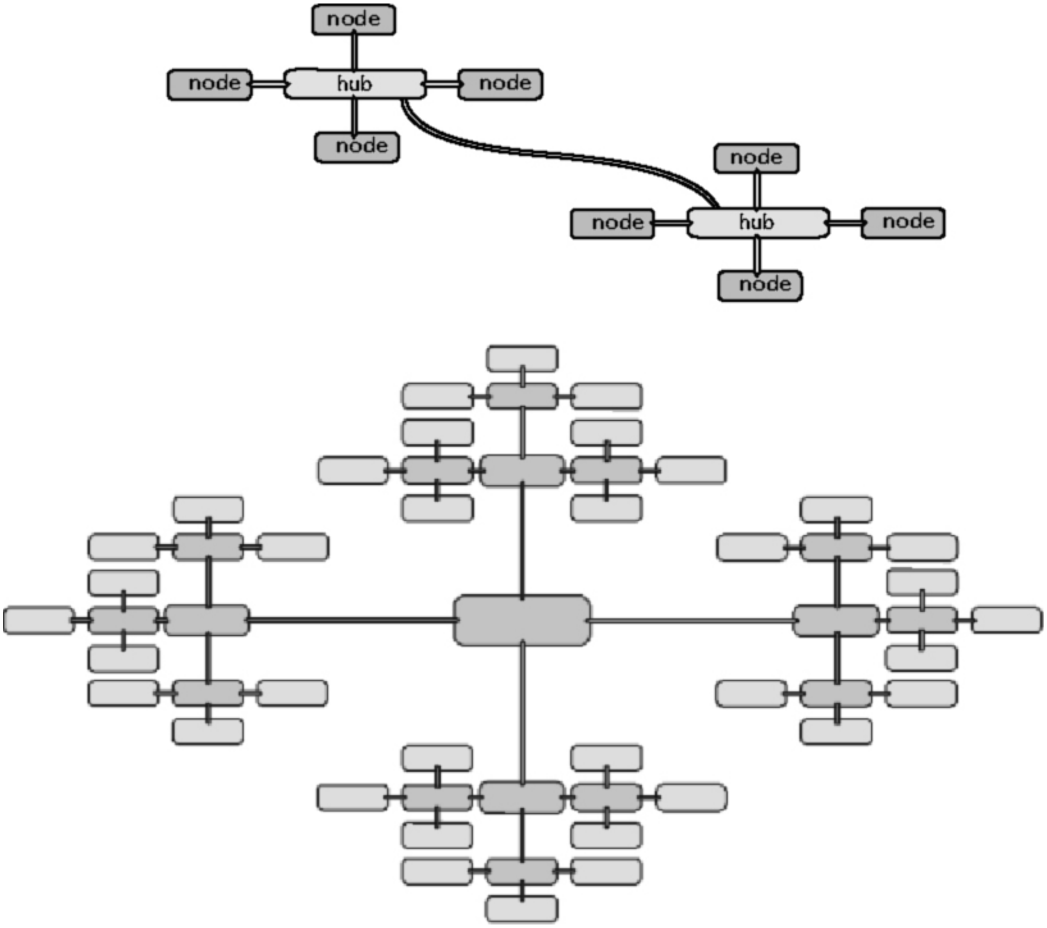


Figure 39-27. Ethernet star of stars topology examples.

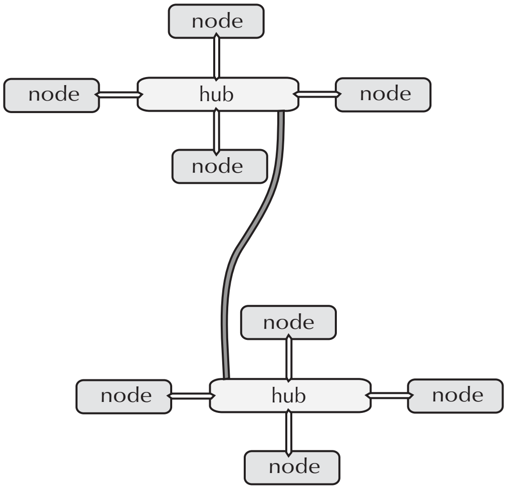


Figure 39-28. Fiber used to interconnect two stars. This allows breaking the copper cable 100 m limitation. Now the only limitation is network diameter. As you can see from Figs. 39-24, 39-25, and 39-26, quite large networks can be made with simple wiring using the star of stars approach. In all these examples, there are no loops. This is because a loop in Ethernet, unless special techniques are used, results in something called a *broadcast storm*, which is sort of the data equivalent of runaway feedback in an audio system.

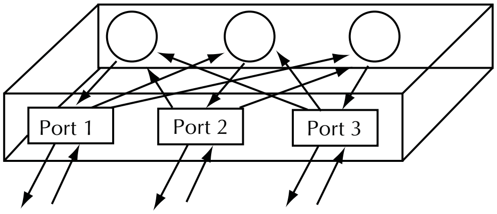


Figure 39-29. Repeater hub functional diagram.

Ethernet switches receive a packet into memory. They examine the destination address and decide which of their ports has attached to it the station with the address in question. Then if the destination is not on the same port as the packet was received from, the switch forwards the packet to only the correct port. Packets where the destination address is on the same port as the packet was received from are discarded.

Switches determine which addresses are connected to each port by recording the source address of every packet received, and associating that address with the port from which it was received. This information is assembled in a look up table (LUT) in the switch. Then as each packet is received, the switch checks to see if the destination address is in the LUT. If it is, the switch knows where to send that packet and only sends it out the appropriate port. If a destination address is not in the LUT, the switch sends that packet out every port but the one from which it was received. Since most Ethernet stations respond to packets received addressed to them, when the response is sent the switch learns which port that address is on. Thereafter packets addressed to that station are only sent out the correct port.

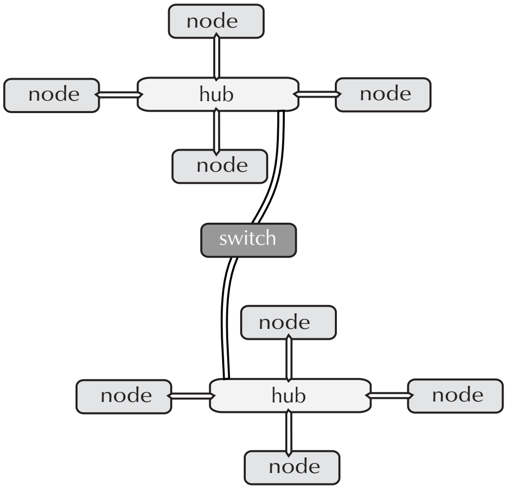


Figure 39-30. An Ethernet switch being used to isolate two collision domains.

If a given MAC address is found on a different port than was contained in the LUT, the LUT is corrected. If no packet is received from a given MAC address within a timeout window of perhaps 5 minutes, its entry in the LUT is deleted. These characteristics allow the switch to adapt and learn as network changes are made.

Packets intended to go out a given port are never allowed to collide inside the switch. Instead each outgoing packet is stored in a first in first out (FIFO) buffer memory assigned to a given port, and transmitted one at a time out the port.

While most data passing through a switch behaves as described above, there is one type of packet that does not. Most data packets are addressed to a specific destination MAC address. This is called unicast addressing. There is a specific address called the multicast or broadcast address. Packets with this address in their destination field are sent to all stations. Therefore, these packets are sent out all ports of a switch except the port they came in on.

Switches are not the shared media of the early coaxial cable Ethernet varieties, or the newer repeater hubs. Instead by storing the packets, examining the addresses, selectively passing the packets on, and FIFO buffering the outputs, they break the network diameter limitation.

Switches have another difference from repeater hubs. Repeater hubs and stations connected to them operate in half duplex mode. In other words a given station can only receive or transmit at different times. If a station that is transmitting in half duplex mode sees a received signal, that tells it a collision has occurred. Since switches store and buffer the packets, they can operate in full duplex mode with other switches or with stations which can operate full duplex. When a station is connected to a switch in full duplex mode it can receive at the same time as it transmits and know that a collision can't occur since it is talking to a full duplex device which does not allow collisions to occur internally, Fig. 39-31.

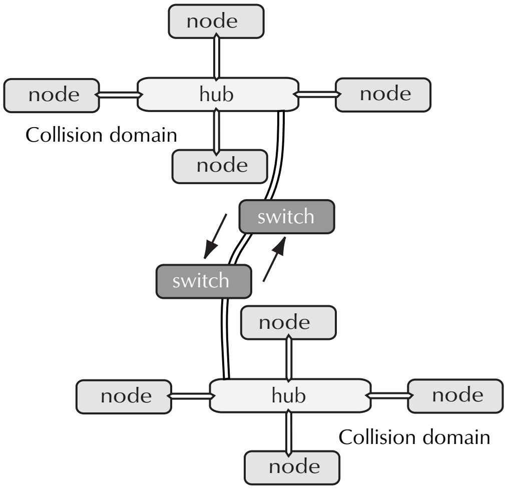


Figure 39-31. Two Ethernet switches showing full duplex operation between them, and the isolation of two collision domains.

Full duplex operation has the added benefit of doubling the communications bandwidth over a half duplex link. A half duplex fast Ethernet connection has 100 MBit/s of available bandwidth which must be split and shared between the packets going each direction on that link. This is because if packets were going both directions at once, that by definition would be a collision. A full duplex link on the other hand has no problem allowing packets to flow in both directions at once, so a fast Ethernet link has 100 MBit/s capability in each direction.

The internal packet routing function inside a switch is called the switch fabric or switch cloud. Switches which contain enough packet routing capability in their cloud to never run out, even if all ports are receiving the maximum possible amount of data, are known as "nonblocking" switches.

Proper Ethernet network design includes ensuring that no packet may go through more than 7 switches on its way from the source to the destination.

When switches were first introduced their expense limited their application to the few situations which required their capabilities. Today the price of switches has come down until they are hardly any more expensive than repeater hubs. As a result the repeater hub is becoming a vanishing part of Ethernet history, Fig. 39-32.

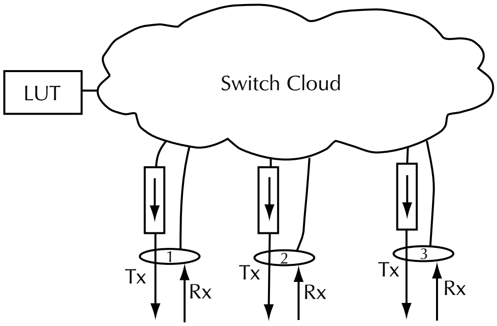


Figure 39-32. Switch functional diagram.

The device that provides this function between a pair of collision domains is called a bridge. As the technology became cheaper multiport bridges started to appear which were called switches. As switches became popular and bridges faded from use, you will sometimes see a bridge referred to as a two-port switch.

Ethernet switches receive a packet into memory. They examine the destination address and decide which of their ports has attached to it the station with the address in question. Then if the destination is not on the same port as the packet was received from, the switch forwards the packet to only the correct port. Packets where the destination address is on the same port the packets were received from are discarded.

Switches determine which addresses are connected to each port by recording the source address of every packet received, and associating that address with the port from which it was received. This information is assembled in a look up table (LUT) in the switch. Then as each packet is received, the switch checks to see if the destination address is in the LUT. If it is, the switch knows where to send that packet and only sends it out the appropriate port. If a destination address is not in the LUT, the switch sends that packet out every port but the one from which it was received. Since most Ethernet stations respond to packets received addressed to them, when the response is sent the switch learns which port that address is on. Thereafter packets addressed to that station are only sent out the correct port.

If a given MAC address is found on a different port than was contained in the LUT, the LUT is corrected. If no packet is received from a given MAC address within a timeout window of perhaps 5 minutes, its entry in the LUT is deleted. These characteristics allow the switch to adapt and learn as network changes are made.

Packets intended to go out a given port are never allowed to collide inside the switch. Instead each outgoing packet is stored in a first in first out (FIFO) buffer memory assigned to a given port, and transmitted one at a time out the port.

While most data passing through a switch behaves as described above, there is one type of packet that does not. Most data packets are addressed to a specific destination MAC address. This is called *unicast addressing*. There is a specific address called the multicast or broadcast address. Packets with this address in their destination field are sent to all stations. Therefore, these packets are sent out all ports of a switch except the port they came in on.

Switches are not the shared media of the early coaxial cable Ethernet varieties, or the newer repeater hubs. Instead by storing the packets, examining the addresses, selectively passing the packets on, and FIFO buffering the outputs, they break the network diameter limitation.

Switches have another difference from repeater hubs. Repeater hubs and stations connected to them operate in half duplex mode. In other words a given station can only receive or transmit at different times. If a station that is transmitting in half duplex mode sees a received signal, that tells it a collision has occurred. Since switches store and buffer the packets, they can operate in full duplex mode with other switches or with stations that can operate full duplex. When a station is connected to a switch in full duplex mode it can receive at the same time as it transmits and know that a collision can't occur since it is talking to a full duplex device that does not allow collisions to occur internally.

Full duplex operation has the added benefit of doubling the communications bandwidth over a half duplex link. A half duplex fast Ethernet connection has 100 MBit/s of available bandwidth that must be split and shared between the packets going each direction on that link. This is because if packets were going both directions at once, that, by definition, would be a collision. A full duplex link, on the other hand, has no problem allowing packets to flow in both directions at once, so a fast Ethernet link has 100 MBit/s capability in each direction.

Of course a repeater hub-based fast Ethernet network has only 100 MBit/s of total bandwidth available for the entire network since it uses a shared media. A network based entirely on fast Ethernet switches has 100 MBit/s of bandwidth available in each direction on each link that makes up the network, assuming that all the stations are capable of full duplex operation. When you combine no collisions with full duplex operation, a switched network can run much faster than a repeater hub-based network.

The internal packet routing function inside a switch is called the *switch fabric* or *switch cloud*. Switches that contain enough packet routing capability in their cloud to never run out of internal bandwidth, even if all ports are receiving the maximum possible amount of data, are known as *nonblocking* switches.

Proper Ethernet network design includes ensuring that no packet may go through more than seven switches on its way from the source to the destination.

When switches were first introduced their expense limited their application to the few situations that required their capabilities. Today the price of switches has come down until they are hardly any more expensive than repeater hubs. As a result the repeater hub is becoming a vanishing part of Ethernet history.

39.11.7 Ethernet Connection Negotiation

It is important to understand how different Ethernet devices negotiate connections between themselves in order to understand why some combinations of devices will work and others won't.

If a 10 MBit/s Ethernet device is not transmitting data, its output stops. After a period of no data transmissions, it will begin sending normal link pulses (NLPs). These allow the device at the other end of the link to know that the connection is still good, and it serves to identify the device as a 10 MBit/s device.

100 MBit/s devices on the other hand always send a signal even when no data is being transmitted. This signal is called a carrier, and serves to identify the device as a 100 MBit/s device.

10/100 Ethernet devices often use a technique called autonegotiation to establish the capabilities of the device at the other end of the link, before the link is established. This process determines if the other device is capable of full or half duplex operation, and if it can connect at 10 MBit/s, 100 MBit/s, or Gigabit speeds. Data is conveyed using fast link pulses (FLPs), which are merely sequences of NLPs that form a message.

If at the end of the autonegotiation process a 10 MBit/s link is established, both devices will send NLPs when idle. If a 100 MBit connection was established, both devices transmit carrier signals.

Autonegotiating devices also utilize parallel detection. This enables a link to be established with a nonnegotiating fixed speed device, and for a link to be established before a device detects the FLPs. The state diagram in Fig. 39-33 shows how the different possible end conditions are reached. Notice that a 100 MBit device can never parallel detect into a full duplex link.

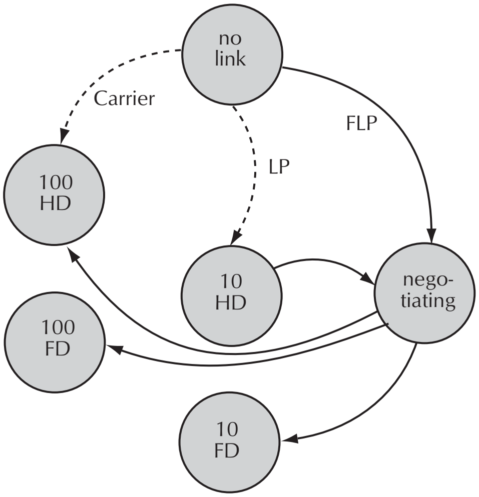


Figure 39-33. Ethernet autonegotiation state diagram. 10 = 10 MBit/s, 100 = 100 MBit/s, HD = half duplex, FD = full duplex. The dashed lines show the parallel detection that can take place while waiting for the FLPs to be recognized. Note that if parallel detection sets up a 100 MBit half duplex connection, then a full duplex connection can never be established.

Fiber optic links do not pass the FLPs needed for autonegotiation though they do pass the carrier. One consequence of this is that if full duplex operation over a fiber link is desired, either manual configuration is required, or an intelligent media converter is required. Such a converter includes circuitry to autonegotiate the link at each end of the fiber.

If a 10/100 NIC were to connect to a 10 MBit/s repeater hub, the 10 MBit/s hub sends NLPs that are detected by the NIC. Seeing the NLPs, the 10/100 NIC goes to 10 MBit/s half duplex operation, and link is established. This is correct since hubs are half duplex devices.

If a 10/100 NIC were to connect to a 100 MBit/s repeater hub, the 100 MBit/s hub sends a carrier that is detected by the NIC. Seeing the carrier, the 10/100 NIC goes to 100 MBit/s half duplex operation, and link is established. This is correct since hubs are half duplex devices.

If a 10/100 NIC were to connect to a 10/100 MBit/s switch, the switch sends FLPs that are detected by the NIC. After interpreting the FLPs, the 10/100 NIC goes to 100 MBit/s full duplex operation, and link is established. This is correct since 100 MBit/s is the highest common rate, and switches can operate full duplex.

A media converter can be thought of as a two-port repeater that coverts from one type of media to another. The most common such conversion is from copper to fiber. There are two basic types of media converters, simple and intelligent.

Simple media converters have no intelligence and just convert electrical signals to light and back. These simple media converters can't pass or detect FLPs. Therefore, they can't pass the signals needed for autonegotiation from one end to the other, nor are they capable of autonegotiating with the ports at each end on their own.

Intelligent media converters add electronics at each end of the fiber link that are able to generate and interpret FLPs. As a result such a converter can either autonegotiate with the port at each end, or be manually configured. They are also capable of both half and full duplex operation.

39.11.8 Managed Switches

All switches have the capabilities described so far. Some switches add significant additional capabilities. Switches with just the basic capabilities are known as *unmanaged switches*, and have become very inexpensive. Unmanaged switches operate automatically, and do not require special settings for their operation. Managed switches provide the capability to control the switch's internal settings.

Common control techniques include a dedicated serial port for control, and Telnet, Web access using a normal Web browser, or Simple Network Management Protocol (SNMP). The last three control methods function over the network. Some managed switches provide all four methods of control. The control capabilities available using each method will often differ. The methods that work over the network will usually require that first the switch is accessed via the serial port, and an IP number assigned to the switch. After that the other control methods can reach the switch over the network at the assigned IP address.

Not all managed switches will provide all the additional capabilities that will be mentioned. Check with the manufacturer of the switch to determine its exact capabilities.

39.11.9 Virtual Local Area Network (VLAN)

Virtual Local Area Network (VLAN) capability allows certain switch ports to be isolated from the other ports. This allows dividing up a larger switch into several virtual smaller switches. While this capability may not matter if all the data is unicast, if there is any multicast or broadcast traffic, there might be significant benefit to isolating that traffic to just certain ports. Some switches may allow data from several VLANs to share a common link to another switch without the data being mingled. Both switches must support the same method for doing this for it to work. Most today use a technique called *tagging* to allow isolated VLANs to share a common physical link between switches.

39.11.10 Quality of Service (QoS)

Quality of service (QoS) allows priority to be given to certain data trying to leave a switch over other data. For example, if we are sending audio over Ethernet using CobraNet®, we would not want there to be any dropouts in the audio if there was a momentary spike in normal computer data traffic through the switch. Such a dropout could occur if a surge in computer data traffic took up bandwidth needed for audio transmission and delayed the reception of the audio packets.

Several means can be used to specify to the switch which traffic is to be given priority. Priority can be given to traffic on a certain VLAN, or that received from certain ports, or that received from certain MAC addresses, or even traffic containing a specific protocol identifier.

39.11.11 Routing

Ethernet switches normally don't examine the payload portion of the Ethernet packet. Routers are capable of looking inside the payload and routing Ethernet packets based on Internet Protocol (IP) addresses that might be found inside some Ethernet payloads. Such a router, or a routing function built into some switches, can allow packets to flow between normally independent networks or VLANs. This can be very useful, for example, to allow a central SNMP management system to monitor and control all the network devices in a facility even if they are in independent isolated networks or VLANs.

39.11.12 Fault Tolerance

39.11.12.1 Trunking

Trunking or link aggregation allows two or more links between switches to be combined to increase the bandwidth between the switches. Both switches must support trunking for this to work. While the algorithm used to share the traffic between the links works for many types of data, it does not for all possible types of data. You may find situations where adding a second link and activating trunking between two switches does not provide any significant increase in available bandwidth, Fig. 39-34.

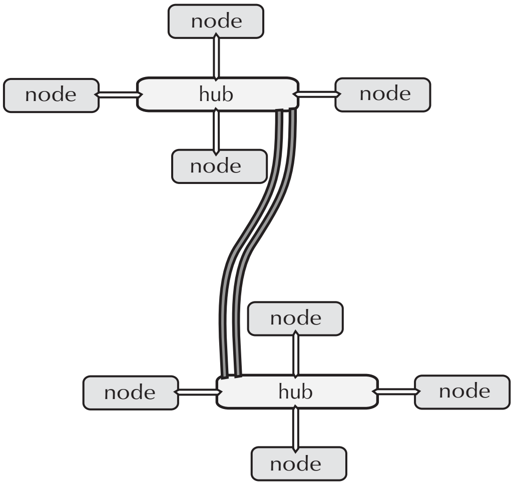


Figure 39-34. Example of trunking between two switches.

Trunking does provide increased fault tolerance, particularly if the links aggregated run through different physical paths between the two switches. If one link is lost, the other link or links will continue to carry traffic between the switches.

39.11.12.2 Spanning Tree

Spanning tree provides automatic protection from inadvertent loops in a network's topology. The cost for this protection is a delay in the activation of a connection made to a port where spanning tree is activated while the switch tries to determine if there is a path out that port that eventually returns to the same switch. If it finds no such path, it will activate the port. The delay might be on the order of 30 seconds to a minute. If it finds a path back to itself on the port, it will disable that port. Whenever a connection to a port is made or lost, and the port has spanning tree active, the switch will reexamine all the ports for loops and activate those where loops are not found.

In any network, damage to the cabling is one of the more common causes of network failures. Spanning tree can be used to reduce the impact of such failures on network operation, Fig. 39-35.

When managed switches with spanning tree capability are used it is common to deliberately build the network with loops. The switches will find the loops and disable enough of the links between switches to insure the network topology is a star of stars and stable. If one of the active links is later disabled, perhaps due to physical damage to the cable or the failure of another switch, then one or more of the currently disabled links will automatically be restored to operation. This provides an inexpensive way to increase the reliability of a network.

Fig. 39-36 shows one possible network topology that can be stable if spanning tree is used. Such a network design can be quite robust, and accommodate multiple failures while maintaining operation.

One difficulty with designing a network that uses spanning tree is that we can't know which links will be disabled and which will stay active. This makes it difficult to predict the amount of traffic a given link will carry.

While spanning tree used with the correct network topology can increase system reliability, it does not respond instantly to failures or changes in network topology. At times it may take several minutes for operation to be restored after a failure.

39.11.12.3 Meshing

At this time meshing is only available on some Hewlett Packard (HP) Procurve switches. Meshing is an attempt to combine the best portions of trunking and spanning tree into a new protocol. Unlike spanning tree, meshing does not disable any links. Instead it keeps track of packets and prevents them from being recirculated around loops. When there are multiple possible routes for a packet to take to its destination, meshing attempts to send the packet by the most direct route.

One of the most significant advantages of meshing is that recovery from failures of links or switches is far faster than spanning tree, and may be accomplished in seconds rather than minutes.

39.11.12.4 Rapid Spanning Tree

More recently a new protocol has brought much of the advantages of meshing and other proprietary technologies into the general market. It allows restoration of a network typically in seconds rather than minutes.

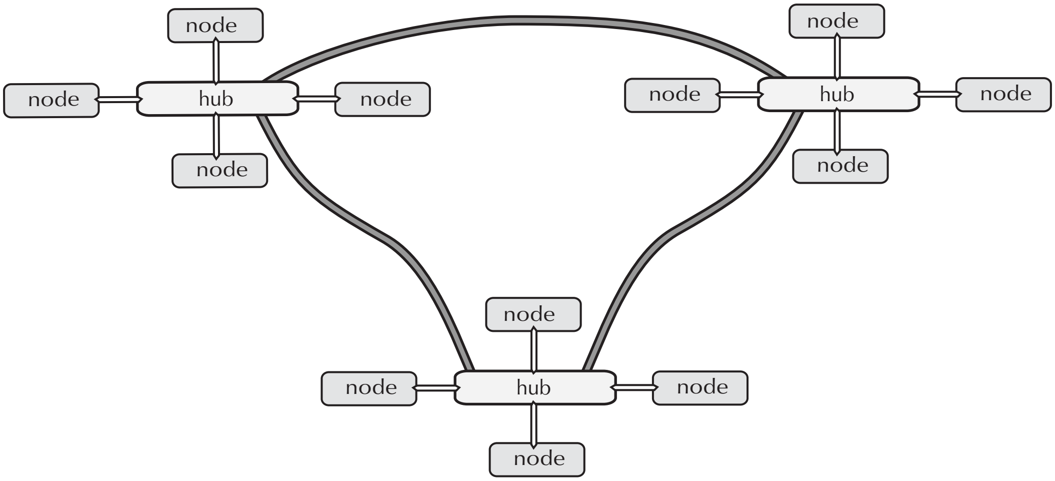


Figure 39-35. A loop around three switches. Spanning tree would disable one of the three links between the switches, and allow the network to be stable.

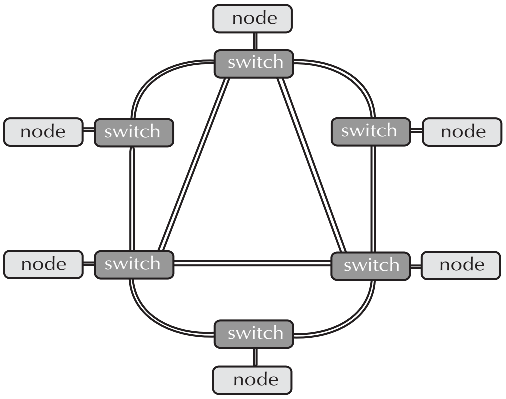


Figure 39-36. Multiple loops around six switches. This is a typical topology used to increase the fault tolerance of a network. Spanning tree would disable enough of the links between the switches to allow the network to be stable.

39.11.13 Core Switching

At its simplest a core switch can be thought of as the central switch in a star of stars configuration. Data that needs only to travel between devices local to each other is switched through the edge switches and never goes through the core switch. Core switches often run at ten times the data rate of the edge switches. For example if the edge switches are fast Ethernet, they will each have a gigabit uplink port that connects back to the gigabit Ethernet core switch.

Besides allowing ten times the data traffic, another reason to use the next higher-speed protocol in the core switch is that the latency through the higher-speed link and switch is only ¹⁄₁₀ as long as if the higher speed was not used.

Some core switches will be equipped with routing capabilities to allow easy central control of all the VLANs using SNMP management.

Core switches are often built to higher-quality standards than ordinary switches since such a core switch can be a single point of failure in the network.

To prevent a single point of failure and greatly increase the fault tolerance of the network, it is possible to use a pair of core switches, each of which connects to all of the edge switches. The network will continue full operation even if one of the core switches or any link to a core switch was to fail.

39.11.14 Ethernet Wiring

Proper design of an Ethernet cable plant is important for reliable operation, ease of maintenance, and maximum performance.

A typical Ethernet network cable path or link is shown in Fig. 39-37. The items that make up the cable plant include:

- Cabling connecting nodes—this can be Cat5 or fiber optic cable.
- Wiring closet patch panels.
- Station cables—the cable that runs from node to wall plate.
- Wall plates—the data or information outlet close to the node.

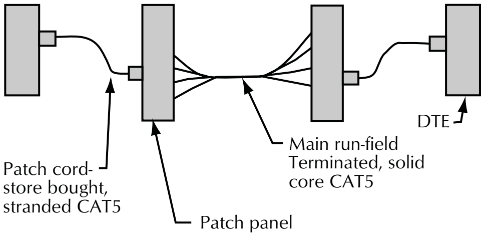

DTE
Main run-field Terminated, solid core CAT5
Patch cord-store bought, stranded CAT5
Patch panel

Figure 39-37. Ethernet typical cable plant showing the entire link from one Ethernet device to another.

It is considered good design practice to include the intermediate patch points as shown. This gives the cable plant operator flexibility in accommodating expansion and configuration changes.

There are two main types of cables used in Ethernet networks: Cat5 cable and fiber optic cable. The following sections will describe these cable types, as well as the issues associated with each.

39.11.14.1 UTP Cable Grades

Unshielded twisted pair (UTP) cables are graded in several categories.

- Quad: nontwisted four conductor formerly used for telephone premise wiring.
- Category 1: No performance criteria UTP.
- Category 2: Rated to 1 MHz (old telephone twisted pair).
- Category 3: Rated to 16 MHz (10Base-T and 100Base-T4 Ethernet, current FCC required minimum for telephone).
- Category 4: Rated to 20 MHz (token-ring).
- Category 5: Rated to 100 MHz—now withdrawn as a Standard and replaced by Cat5e.
- Category 5e: Improved Cat5 with tighter tolerances. (100Base-TX and 1000Base-T Ethernet).
- Category 6: Rated to 250 MHz.
- Category 7: Shielded cabling mostly used in Europe.

When used with quality equipment there usually is not a lot of advantage for fast Ethernet networks in using cable with a rating beyond Cat5e. Future higher speed networks, or marginal equipment on fast Ethernet may benefit from improved cable.

Unless specified differently by the manufacturer, most UTP has a minimum bend radius of 4 times the cable diameter or about 1 inch.

Cat5e is inexpensive unshielded twisted pair (UTP) data grade cable. It is very similar to ubiquitous telephone cable but the pairs are more tightly twisted. It should be noted that not all Cat5e cable is UTP. Shielded Cat5 also exists but is rare due to its greater cost and much shorter distance limitations than UTP Cat5e.

39.11.14.2 Distance Limitations

On fast Ethernet systems, Cat5e cable runs are limited to 100 m due to signal radiation and attenuation considerations. A Cat5e run in excess of 100 m may be overly sensitive to electromagnetic interference (EMI).

39.11.15 Connectors

Cat5 cable is terminated with an RJ-45 connector. Strictly speaking this nomenclature is incorrect since it designates a particular telephone usage of the connector rather than the connector itself. Since 8 position 8 contact nonkeyed modular connector is difficult to say and write, we are stuck with the common usage of RJ-45, Fig. 39-38.

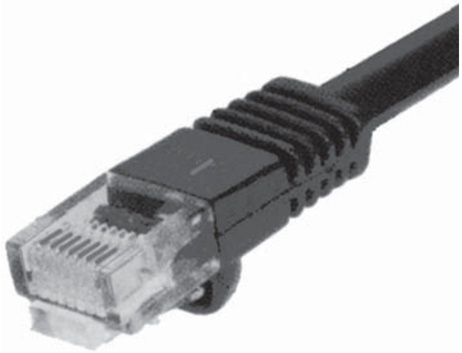


Figure 39-38. "RJ-45" plug. This is an 8 position 8 contact nonkeyed modular connector originally developed for telephone applications. It is the connector used for all UTP cables in Ethernet networks.

There are two different types of contacts in RJ-45 connectors. There is the *bent tyne* contact, intended for use with solid core Cat5, and then there is the *aligned tyne* contact used with stranded Cat5 cable. Errors can occur when using an incorrect cable/connector combination. Fig. 39-39 shows an end-on view of a single contact in a modular connector. The aligned tyne contact (on left) must be able to pierce through the center of the wire, therefore it can only be used on stranded wire. The bent tyne contact has the two or three tynes offset from each other to straddle the conductor; therefore, it can be used on solid or stranded wire.

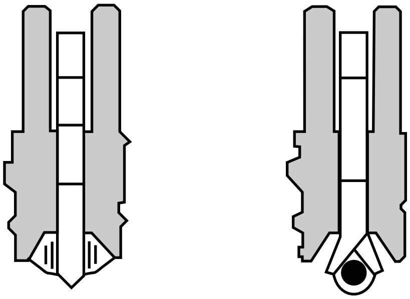


Figure 39-39. Different style contacts found in RJ-45 plugs. On the left is the aligned tyne contact style such as found on original telephone connectors. Since it must pierce through the wire, it must only be used on stranded wire. On the right is the bent tyne contact. It has two or three-tynes which are offset from each other and straddle the conductor. As a result it can be used on either stranded or solid conductor wire.

Cable openings in modular connectors can be shaped for flat, oval, or round cable. Cat5 cable does not usually fit properly into connectors made specifically for flat telephone cable, Fig. 39-40.

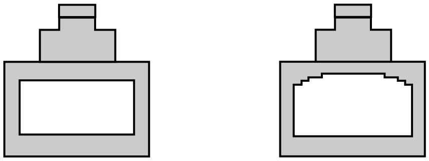


Figure 39-40. Different style cable openings found in RJ-45 plugs. On the left is the opening for flat cable such as found on original telephone connectors. On the right is the opening for round cable needed for Cat 5 cable.

Cheap modular connectors may not have proper gold plating on the contacts, but instead only have a gold flash. Without proper plating, the connectors may quickly wear and corrode, causing unreliable connections.

AMP makes quality modular connectors, but the secondary crimp point is located in a different position from everyone else's connectors. Fig. 39-41 shows a standard crimp die and an AMP plug. Point "A" is the primary crimp point, and should fold the primary strain relief tab in the plug down so that it locks against the cable jacket. At the opposite end of the plug, the contacts are pressed down into the individual conductors. The "B" secondary crimp point secures the individual conductors so that they do not pull out of the contacts.

AMP puts this crimp in a different location from all other manufacturers. If AMP connectors are used in a standard crimper they will either jam, bend, or break the crimp die. If standard connectors are used in an AMP crimper, the die will usually break. Once either type of plug is properly crimped onto the wire, they are interchangeable and will work properly in any mating jack, Fig. 39-41.

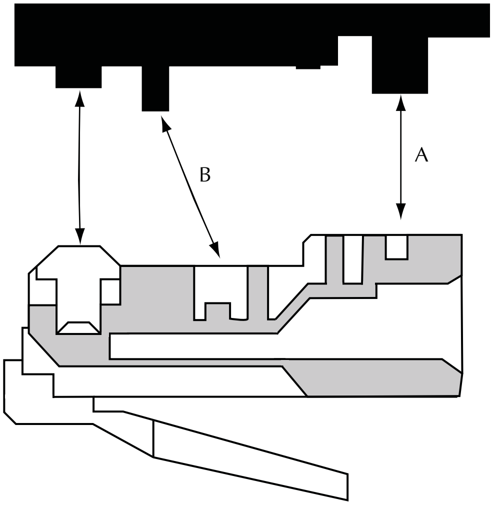


Figure 39-41. A standard crimp die above an AMP connector. Note the misalignment of the secondary strain relief crimp point B. An AMP connector will jam in a standard crimper, while a standard RJ-45 will usually break an AMP crimp die.

Some plugs are made with inserts that guide the wires. These can make the job of properly assembling the connector easier. Some connectors made with inserts may also provide better performance than Cat5, Figs. 39-42 and 39-43.

39.11.15.1 Pairing, Color Codes, and, Terminations

Cat5 cable consists of four twisted pairs of wires. To minimize the crosstalk between the pairs, each pair is twisted at a slightly different rate. For fast Ethernet, one pair is used to transmit (pins 1 and 2) and another pair is used to receive (pins 3 and 6). The remaining two pairs are terminated but unused by fast Ethernet. Although only two of the four twisted pairs are used for fast Ethernet, it is important that *all pairs be terminated*, and that the proper wires be twisted together. Standards set forth by EIA/TIA 568A/568B and AT&T 258A define the acceptable wiring and color-coding schemes for Cat5 cables. These are different from the USOC wiring Standards used in telecommunications, Figs. 39-44 and 39-45.

Note that there are two conflicting color code standards for data use of the RJ-45 connector. Both work just fine, but to avoid problems make sure that one of the standards is selected and used uniformly throughout a facility.

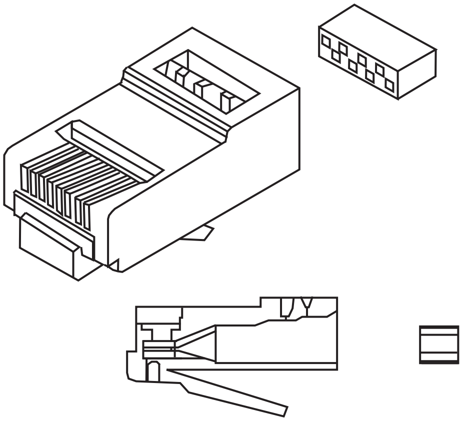


Figure 39-42. An RJ-45 connector that uses an insert to guide the individual conductors into the connector. Such a connector can be easier to properly assemble on a cable. Some connectors made this way provide performance well beyond Cat5.

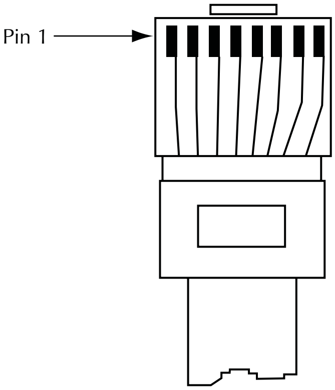
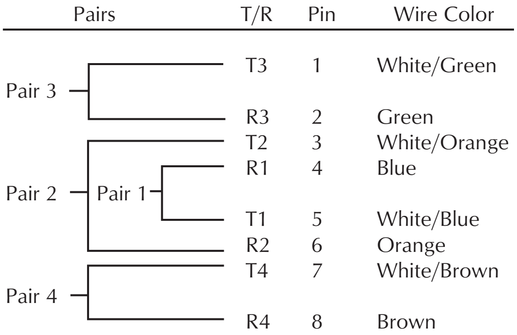


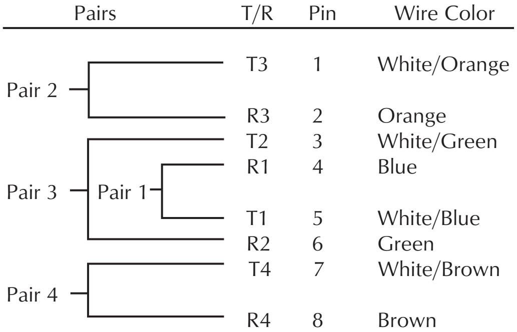
Figure 39-43. RJ-45 pin numbering.

Often when installers accustomed to telephone wiring install data cabling they will incorrectly use the telephone USOC (Universal Service Ordering Code) wiring scheme. This will result in a network that either does not work, or has very high error rates, Fig. 39-46.



Pairs	T/R	Pin	Wire Color
Pair 3	T3	1	White/Green
Pair 3	R3	2	Green
Pair 2	T2	3	White/Orange
Pair 1	R1	4	Blue
Pair 1	T1	5	White/Blue
Pair 2	R2	6	Orange
Pair 4	T4	7	White/Brown
Pair 4	R4	8	Brown

Figure 39-44. Standard EIA/TIA T568A (also called ISDN, previously called EIA). One of the wiring schemes used for Ethernet.

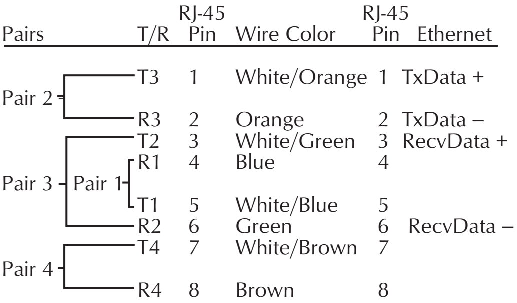


Pairs	T/R	Pin	Wire Color
Pair 2	T3	1	White/Orange
Pair 2	R3	2	Orange
Pair 3	T2	3	White/Green
Pair 1	R1	4	Blue
Pair 1	T1	5	White/Blue
Pair 3	R2	6	Green
Pair 4	T4	7	White/Brown
Pair 4	R4	8	Brown

Figure 39-45. Standard EIA/TIA T568B (also called AT&T specification, previously called 258A). One of the wiring schemes used for Ethernet.

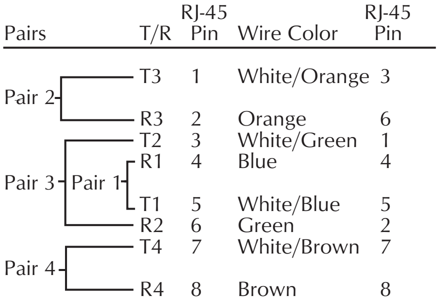
Figure 39-46. USOC (Universal Service Order Code), the wiring scheme used for telephones. This must not be used for data! The eight contact connector uses all pairs for four lines. The six contact connector uses only center one–three pairs for one, two, or three line phones.

Normal Ethernet cable wiring such as shown in Fig. 39-47 is used for interconnecting unlike devices. In other words it is used to connect the Network Interface Card (NIC) in a station to a switch or repeater hub. Connections between like devices such as a pair of NICs, or between switches, repeater hubs, or switch to repeater hub require a "crossover" cable wired per Fig. 38-48. This is because the data transmit pair must connect to the receive input, and vice versa.



Pairs	T/R	RJ-45 Pin	Wire Color	RJ-45 Pin	Ethernet
Pair 2	T3	1	White/Orange	1	TxData +
Pair 2	R3	2	Orange	2	TxData –
Pair 3	T2	3	White/Green	3	RecvData +
Pair 1	R1	4	Blue	4	
Pair 1	T1	5	White/Blue	5	
Pair 3	R2	6	Green	6	RecvData –
Pair 4	T4	7	White/Brown	7	
Pair 4	R4	8	Brown	8	

Figure 39-47. Ethernet Standard (T568B colors) patch cord wiring used for most interconnects. Ethernet usage shown for 10Base-T and 100Base-T. Gigabit Ethernet uses all the pairs.



Pairs	T/R	RJ-45 Pin	Wire Color	RJ-45 Pin
Pair 2	T3	1	White/Orange	3
Pair 2	R3	2	Orange	6
Pair 3	T2	3	White/Green	1
Pair 1	R1	4	Blue	4
Pair 1	T1	5	White/Blue	5
Pair 3	R2	6	Green	2
Pair 4	T4	7	White/Brown	7
Pair 4	R4	8	Brown	8

Figure 39-48. Ethernet standard (T568B colors) crossover cord. Pairs 2 and 3 are reversed end to end. Used for connections between like devices (NICs, switches or repeater hubs).

It is very easy to tell the difference between a crossover cable and a straight-through cable by looking at the conductors in the RJ-45 connectors. If the wiring is identical at both ends, you are holding a straight-through cable, if it is different, you most likely have a crossover cable.

Some hubs and switches have uplink ports that can eliminate the need for crossover cables. Such a port is wired with pairs 2 and 3 reversed internally. Make sure that when connecting two switches or repeater hubs so equipped, you only use the uplink port at one end. Another caution is that often such an uplink port is not an independent port but is wired internally to one of the normal ports. In such a case make sure that only one of the pair of ports is used.

Some switches employ an autoselect crossover feature. This allows the use of either a straight-through or a crossover cable on any port. The switch automatically senses which cable type is in use and adjusts the electronics to suit the cable.

Stranded patch cable sometimes has different colors.

Pair 1 Green and Red

Pair 2 Yellow and Black
Pair 3 Blue and Grey
Pair 4 Brown and Grey

39.11.16 Fiber Optic Cable

There are two basic varieties of fiber optic cable, single-mode and multimode. Both are used in Ethernet network designs. Two fibers are needed to make an Ethernet connection, one fiber for transmit, and one for receive, Fig. 39-49.

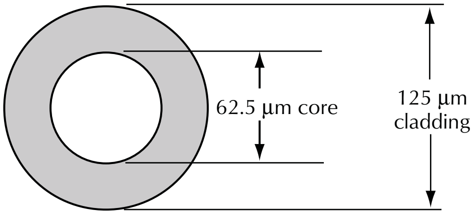


Figure 39-49. Cross-section through a multimode fiber optic cable.

Multimode fiber is built of two types of glass arranged in a concentric manner. Multimode fiber allows many modes, or paths, of light to propagate down the fiber optic path. The relatively large core of a multimode fiber allows good coupling from inexpensive LED light sources, and the use of inexpensive couplers and connectors, Fig. 39-50.

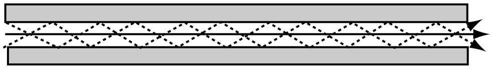


Figure 39-50. Possible light paths down a multimode fiber optic cable. Note that there are multiple possible paths for the light to take. This is why this is called multimode cable.

Two sizes of multimode fiber are available. 62.5/125 µm is used primarily in data communications, and 50/100 µm is used primarily in telecommunications applications. The standard for transmission of 100 Mbit Ethernet over 62.5/125 µm multimode fiber is called 100Base-FX. 100Base-FX has a 2 km distance limitation.

Single-mode fiber optic cable is built from a single type of glass. The cores range from 8 µm to 10 µm, with 8/125 µm being the most commonly used. There is only a single path of light through the fiber, Fig.39-51.

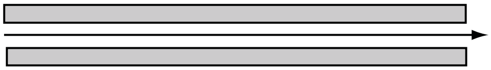


Figure 39-51. Possible light paths down a single-mode fiber optic cable. Note that there is only one possible path for the light to take. This is why this is called single-mode cable.

While single-mode fiber cable costs approximately the same as a multimode cable, the cost of the optical transmitters and receivers is significantly more for a single-mode installation than multimode. Single-mode fiber has a core diameter that is so small that only a single mode of light is propagated. This eliminates the main limitation to bandwidth, but makes coupling light into the fiber more difficult.

Although multimode fiber cable has a specific distance limitation of 2 km, distance limitations of single-mode fiber vary according to the proprietary system in use. All are in excess of 2 km with some allowing 100 km. There is currently no Ethernet standard for single-mode fiber.

39.11.16.1 Fiber Optic Connectors

There are two common types of fiber optic connectors, SC and ST, Fig. 39-52. The ST, or straight tip, connector is the most common connector used with fiber optic cable, although this is no longer the case for use with Ethernet. It is barrel shaped, similar to a BNC connector, and was developed by AT&T. A newer connector, the SC, is becoming more and more popular. It has a squared face and is thought to be easier to connect in a confined space. The SC is the connector type found on most Ethernet switch fiber modules and is the connector of choice for 100 Mbit and gigabit Ethernet. A duplex version of the SC connector is also available, which is keyed to prevent the TX and RX fibers being incorrectly connected.

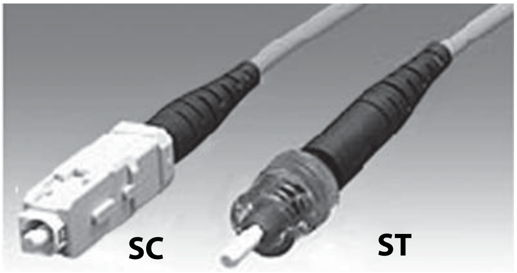


Figure 39-52. The two most common fiber optic connectors. The ST on the right has been the most popular. Today most Ethernet fiber optic interfaces come equipped for the SC on the left.

There are two more fiber connectors that we may see more of in the future. These are the MTRJ and MTP. They are both duplex connectors and are approximately the size of an RJ-45 connector.

39.11.16.2 Cabling and Network Performance

A number of factors can degrade the performance of an Ethernet network, and among these is a poor cable plant. Cabling problems and a susceptibility to EMI can actually lead to packet loss. The following sections present cabling considerations that will help to insure a high-quality cable plant installation.

The Cat5 specifications require that no more than ½ inch of the pairs be untwisted at each termination. It is good practice to never strip more of the outer jacket of the cable than is required, and to keep the cable pairs twisted at their factory twist rates until the point they must separate to enter the terminations.

As with audio cabling, there are certain proximity specifications to be aware of when designing your network cable routes. Fig. 39-53 lists some UTP proximity guidelines. For fiber optic cable runs, proximity is not a concern due to fiber's inherent immunity to EMI and RFI.

Condition	<2kVA	2-5k VA	>5kVA
Unshielded power lines or electrical equipment in proximity to open or nonmetal pathways	5 in (12.7 cm)	12 in (30.5 cm)	24 in (61 cm)
Unshielded power lines or electrical equipment in proximity to grounded metal conduit pathway	2.5 in (6.4 cm)	6 in (15.2 cm)	12 in (30.5 cm)
Power lines enclosed in a grounded metal conduit (or equivalent shielding) in proximity to grounded metal conduit pathway	N/A	6 in (15.2 cm)	12 in (30.5 cm)
Transformers and electric motors	40 in (1.02 m)	40 in (1.02 m)	40 in (1.02 m)
Fluorescent lighting	12 in (30.5 cm)	12 in (30.5 cm)	12 in (30.5 cm)

Figure 39-53. Ethernet UTP proximity specifications. Fiber is insensitive to electromagnetic fields and does not require separation.

39.11.17 Cable Installation

39.11.17.1 Cable Ties and UTP

Another factor that can degrade the installation quality is snug cable ties. Ties should never be pulled tight enough to deform or dent the outer jacket of the UTP cable. Doing so produces a slight change in the cable impedance at the point under the tie, which can lead to poor network performance. If tight ties are used at even intervals down the cable length, the performance degradation is even worse.

For best performance with minimum alien crosstalk between cables, they should not be bundled or combed into a straight and neat harness, but instead be allowed to lie randomly and loosely next to each other.

39.11.17.2 Pull Force and Bend Radius

A common myth is that fiber optic cable is fragile. In fact, an optical fiber has greater tensile strength than copper or steel fibers of the same diameter. It is flexible, bends easily, and resists most of the corrosive elements that attack copper cable. Some optical cables can withstand pulling forces of more than 150 pounds! The fact is, Cat5 cable may be more fragile than optical cables: tight cable ties, excessive untwisting at the connector, and sharp bends can all degrade the cable's performance until it no longer meets Cat5 performance requirements. While fiber may have a reputation for being more fragile than it really is, it still has limitations, and as such, care should be taken when installing both Cat5 and fiber optic cables. Here are some guidelines for Cat5 and fiber optic bend radius and pull force limitations.

39.11.17.3 Cat 5

All UTP cables have pull force limitations much lower than those tolerated in the audio industry. If more than 25 lbs of force is applied to Cat5 cable during installation, it may no longer meet specification. Like most audio cables, UTP cables also have minimum bend radius limitations. Generic Cat5 allows a minimum bend radius of four times the cable diameter or 1 inch for a ¼ inch diameter cable. Unless specified otherwise by the manufacturer, it is fairly safe to use this as a guideline. Note that this is a minimum bend *radius* and not a minimum bend *diameter*.

39.11.17.4 Fiber Optic Cable

The bend radius and pull force limitations of fiber vary greatly based on the type and number of fibers used. If no minimum bend radius is specified, one is usually safe in assuming a minimum radius of ten times the outside diameter of the cable. For pulling force, limitations begin at around 50 lbs and can exceed 150 lbs. In general, it is recommended that you check with the fiber manufacturer for specifications on the specific cable used in your installation.

39.11.17.5 Cable Testing

All network cable infrastructure, both copper and fiber, should be tested prior to use, and after any suspected damage. The tester used should certify the performance of the link as meeting the Cat5, Cat5e, Cat6, or whatever performance level you thought you had bought.

Inexpensive Cat5 testers are often just continuity checkers and are worse than useless since they can provide a false sense of security that the cabling is fine when in fact it may be horrible.

A tester that can correctly certify a link as meeting all of the Cat5 specifications will cost thousands of dollars, and testers capable of certifying to higher levels are more expensive.

While there are dedicated fiber testers, many of the quality Cat5 testers can accept fiber testing modules.

39.12 CobraNet®

CobraNet® is a technology developed by Peak Audio a division of Cirrus Logic, Inc., for distributing real-time, uncompressed, digital audio over Ethernet networks. The basic technology has applications far beyond audio distribution, including video and other real-time signal distribution.

CobraNet® includes specialized Ethernet interface hardware, a communications protocol that allows isochronous operation over Ethernet, and firmware running on the interface that implements the protocol. It can operate on either a switched network or a dedicated repeater network.

To the basic Ethernet capabilities, CobraNet® adds transportation of isochronous data, sample clock generation and distribution, and control and monitoring functions.

The CobraNet® interface performs synchronous to isochronous and isochronous to synchronous conversions as well as the data formatting required for transporting real time digital audio over the network.

A CobraNet® interface provides conversion from synchronous to isochronous and back, and formats the data to meet Ethernet requirements. This allows it to provide real-time digital audio across the network.

As shown in Fig. 39-54, CobraNet® can transport audio data, and carry and use control information as well as allowing normal Ethernet traffic over the same network connection. Simple Network Management Protocol (SNMP) can be used for control and monitoring. In most cases normal Ethernet traffic and CobraNet® traffic can share the same physical network.

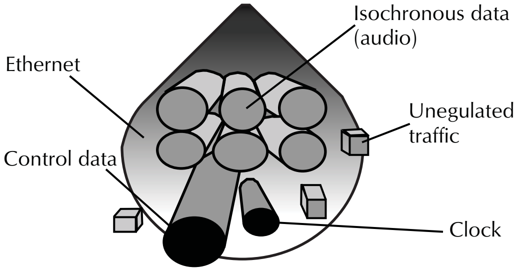


Figure 39-54. CobraNet® Data Services showing the different types of data flowing through the Ethernet network.

39.12.1 CobraNet® Terminology

CobraNet® Interface.

The hardware or hardware design with associated firmware provided by Peak Audio to CobraNet® licensees and affiliates.

CobraNet® Device. A product that contains at least one CobraNet® interface.

Conductor. The particular CobraNet® interface selected to provide the master clock and transmission arbitration for the network. The other CobraNet® interfaces in the network function as performers.

Audio Channel. A 48 kHz sampled digital audio signal of 16, 20 or 24 bit depth.

Bundle. The smallest unit for routing audio across the network. Each bundle is transmitted as a single Ethernet packet every isochronous cycle, and can carry from zero to eight audio channels. Each bundle is numbered in the range from 1 to 65,535. A given bundle can only be transmitted by a single CobraNet® interface. There are two basic types of bundles.

Multicast Bundle. Bundles 1 through 255 are multicast bundles and are sent using the multicast MAC destination address. If a transmitter is set to a multicast bundle number it will always transmit regardless of whether a receiver is set to the same bundle number. Multiple receivers can all pick up a multicast bundle.

Unicast Bundle. Bundles 256 through 65,279 are unicast bundles and are sent using the specific MAC destination address of the receiver set to the same bundle number. Only a single receiver can receive each of these bundles. If no receiver is set for this bundle number the bundle will not be transmitted.

39.12.2 Protocol

CobraNet® operates at the data link layer (OSI Level 2). It uses three distinct packet types, all of which are identified in the Ethernet packet by the unique protocol identifier (8819 hex) assigned by the IEEE to Peak Audio. CobraNet® is a local area network (LAN) technology and does not utilize Internet Protocol (IP), which is most important in wide area networks (WAN).

39.12.3 Beat Packet

Beat packets are sent with the multicast destination MAC address 01:60:2B:FF:FF:00. They contain the clock, network operating parameters, and transmission permissions. The beat packet is sent by the Conductor and indicates the start of the isochronous cycle. Because the beat packet carries the clock for the network, it is sensitive to delay variations in its delivery to all the other CobraNet® interfaces. Failure to meet the delay variation specification can keep the other devices from locking their local clocks to the master clock in the Conductor. The beat packet is usually small on the order of 100 bytes, but grows with the number of active bundles.

39.12.4 Isochronous Data Packet

One isochronous data packet is transmitted for each bundle each isochronous cycle, and carries the audio data. It can be addressed to either unicast or multicast destination addresses depending on the bundle number. Since the CobraNet® interfaces buffer the data, out of order delivery within an isochronous cycle is acceptable. To reduce the impact of the Ethernet packet structure overhead on the total bandwidth consumed, data packets are usually large on the order of 1000 bytes.

39.12.5 Reservation Packet

Reservation packets are sent with the multicast destination MAC address 01:60:2B:FF:FF:01. CobraNet® devices usually send a reservation packet once per second. This packet is never large.

39.12.6 Timing and Performance

In order for CobraNet® to provide real-time audio delivery, certain maximum delay and delay variation requirements must be put on the performance of the Ethernet network, Fig. 39-55.

If the network loses a beat packet it will cause an interruption in proper operation of the entire CobraNet® network. If an isochronous data packet is lost, a 1⅓ ms dropout will occur only in the audio carried by that particular bundle. A single such dropout may be inaudible or may make a “tick” in the audio. Large numbers of dropouts may sound like distortion.

Parameter	Minimum	Maximum	Typical	Comments
Isochronous Cycle Interval			1333 µs	Future CobraNet® revisions may allow other cycle interval options.
Beat Packet Length	5.12 µs	121.4 µs	≤10 µs	Beat packet grows as bundle count increases.
Data Packet Length	5.12 µs	121.4 µs	100 µs	Size dependent on audio resolution and number of audio channels carried in bundle.
Reservation Packet Length	5.12 µs	121.4 µs	10 µs	
Inter-Packet Spacing	0.96 µs		5 µs	
Beat Packet Delay Variation	0 µs	250 µs		Normal delay distribution assumed.
Forwarding Delay	0 µs	400 µs		Assume maximal packet length when calculating store and forward delay (if applicable). Includes delay variation, i.e., 750 µs forwarding delay + 250 µs maximal positive excursion due to delay variation = 1000 µs. A higher forwarding delay can be tolerated on networks with small delay variation. If the forwarding delay specification is exceeded, additional delay is automatically added to the audio in increments of 64 sample periods (1⅓ ms).

Figure 39-55. CobraNet® packet timing and performance requirements for the Ethernet network. Make sure your Ethernet switch vendor will guarantee that their switches in the configuration you propose will meet the above delay variation and forwarding delay specifications.

Hexadecimal Bundle Number	Decimal Bundle Number	Designation	Usage	Transmission Addressing	Transmission Mode
0	0	Null	Unused bundle. Disables transmission/reception when selected.	Never transmitted.	Never transmitted.
1–FF	1–255	Multicast	Publicly available bundles. Each bundle is transmitted by a single unit and may be received by any number of units.	Always multicast.	Always transmitted.
100–FEFF	256–65279	Unicast	Publicly available bundles. Each bundle is transmitted by a single unit. If the default unicast mode setting is used, it will only be received by a single unit.	Generally unicast, but may multicast if *txUnicastMode* variable is so adjusted.	Only transmitted when at least one receiver is identified via reverse reservation.
FF00–FFFF	65280–65535	Private	Individual transmitters locally allocate private bundles. The bundle number is conditioned on the transmitter's MAC. There are 256 of these bundles per transmitter thus the total number of private bundles is virtually unlimited.	Generally unicast, but may multicast if *txUnicastMode* variable is so adjusted.	Only transmitted when at least one receiver is identified via reverse reservation.

Figure 39-56. CobraNet® bundle types. The bundle number specifies the type of bundle. There can only be a single transmitter for any given multicast or unicast bundle number at a time on the same network or VLAN. The three bundle types have different characteristics. Multicast bundles will be routed to every port in a switched network, or to every port in a VLAN, so use them sparingly. It is generally suggested that no more than four multicast bundles be used at any one time in a given switched network or VLAN within a switched network.

39.12.7 Bundle Identification

Audio is carried over CobraNet® networks in bundles. Bundles may contain from zero to eight audio channels. Each bundle consists of a stream of packets of one of three types specified in Fig. 39-56.

39.12.8 Multicast Bundles

In a given network or VLAN there can only be a single instance of a given multicast bundle number at a time. The conductor will only allow one CobraNet® transmitter to be active during any isochronous cycle on a given multicast bundle number.

Multicast bundles are always multicast addressed, and are always transmitted even if no receiver has selected that bundle. Since they are multicast, they will appear on every port of a network or VLAN, even on switched networks. Therefore, the receiver does not have to submit a reverse reservation request to the conductor in order to receive a multicast bundle since the bundle will always appear at its input.

Caution must be used with multicast bundles and switched networks so as not to overwhelm the ports with multicast traffic. It is generally suggested to not use more than four multicast bundles on a given switched network or VLAN at the same time.

Multicast bundles can serve as a common denominator to allow interoperability with CobraNet® devices, which can only be configured from their front panel switches.

39.12.9 Unicast Bundles

In a given network or VLAN there can only be a single instance of a given unicast bundle number at a time. The conductor will only allow one CobraNet® transmitter to be active during any isochronous cycle on a given unicast bundle number.

Unicast bundles may be either unicast or multicast addressed based on the transmitter's reception of one or more reverse reservation requests for its bundle number. The *txUnicastMode* variable is used to control the transmitter's ability to switch to multicast on a unicast bundle number. The default setting of the *txUnicastMode* variable disables the ability to transmit multicast on a unicast bundle number. With the default setting, if more than one receiver requests a given unicast bundle number, only the first receiver to get its reverse reservation request in will get that bundle. With the default setting, unicast bundles can't be used for point to multipoint routing, instead multicast bundles must be used.

Some CobraNet® devices allow the same audio to be transmitted on more than one bundle at a time. This can provide an alternative way for a single CobraNet® device to unicast to as many as four receiving devices at the same time.

Unicast bundles are only transmitted if a receiver has requested that bundle. This allows a receiver to select which of many sources it wishes to receive, and only the source selected will transmit onto the network.

39.12.10 Private Bundles

Individual CobraNet® transmitters® control their own private bundles. Unlike multicast and unicast bundles, there may be more than one private bundle with the same bundle number on a network or VLAN at the same time. This is because a private bundle is specified using the transmitter's unique MAC address in addition to the bundle number.

Private bundles may be either unicast or multicast addressed based on the transmitter's reception of one or more reverse reservation requests for its bundle number and MAC address. The *txUnicastMode* variable is used to control the transmitter's ability to switch to multicast on a private bundle number. The default setting of the *txUnicastMode* variable disables the ability to transmit multicast on a private bundle number. With the default setting, if more than one receiver requests a given private bundle number from the same transmitter, only the first receiver to get its reverse reservation request in will get that bundle. With the default setting, private bundles can't be used for point to multipoint routing, instead multicast bundles must be used.

Some CobraNet® devices allow the same audio to be transmitted on more than one bundle at a time. This can provide an alternative way for a single CobraNet® device to send private bundles to as many as four receiving devices at the same time.

Private bundles are only transmitted if a receiver has requested that bundle. This allows a receiver to select which of many sources it wishes to receive, and only the source selected will transmit onto the network.

39.12.11 Bundle Assignments

Over CobraNet®, all audio channels are packaged into groups called bundles for transmission over the Ethernet network. The usual assignment is eight audio channels of 20 bit depth into one bundle. This is the maximum size possible, although using less audio channels is possible. In general for most efficient utilization of network bandwidth, maximum-size bundles are suggested. In the rest of this section we will be talking about maximum-size bundles. If 24 bit audio channels are used the maximum is seven audio channels packaged into a single bundle due to the maximum allowable Ethernet packet size.

A CobraNet® system is coordinated by one of the devices called the conductor. When two or more CobraNet® devices are interconnected properly, one of the devices will be elected the network conductor based on a priority scheme. The conductor indicator will light on the CobraNet® device that is serving as the conductor.

Each CobraNet® device has the ability to send and receive a fixed number of bundles. The bundle number tells the CobraNet® conductor which specific CobraNet® device is trying to communicate with which other CobraNet® device(s) over the network. Use of bundle numbers removes the necessity of the user having to tell the devices the Ethernet hardware (MAC) addresses of the other devices with which it is trying to communicate. As long as the CobraNet® devices are all set to the same bundle number, the CobraNet® system takes care of all the rest of the technical details of setting up an audio path over Ethernet between the devices.

A given bundle may have only one transmitter that places it onto the network. Unicast bundles may have only a single receiver. Multicast bundles may have multiple receivers.

In an ordinary Ethernet data network it is possible to mix both repeater hubs and switches and have the network continue to work. This is not the case with CobraNet®! For a CobraNet® network, you must either use all repeater hubs, or all switches in the network. This is because the CobraNet® protocol changes depending on which type of network it is operating over. However, non-CobraNet® devices may be attached to a switched CobraNet® network via repeater hubs.

On a repeater hub-based network, there is a fixed maximum of eight bundles per network. Any bundle may be placed onto the network from any port, and will appear at every other port on the network. The bundles usually used in a repeater hub network are numbered in the range from 1 to 255 decimal, and are called multicast bundles. Such bundles are always transmitted in a multicast mode, and may be received by any of the CobraNet® devices on the network.

As long as the limit of eight total bundles is not exceeded, it does not matter which channel numbers in the range of 1 to 65,279 are used.

It is not suggested to mix ordinary computer data on a repeater network with CobraNet®, as this could result in dropouts in the audio.

On a switched network, there is no fixed maximum number of bundles possible. The number will be determined by the network design. Again, bundles from 1 to 255 decimal are multicast bundles and, since they are multicast, will usually be sent to every port in the network. It is not suggested to use more than four multicast bundles in a switched CobraNet® network. There are special cases where more could be used, which we will go into later.

Bundles from 256 to 65,279 decimal are called unicast bundles. These are addressed to a single destination unit and are usually sent unicast. A switch will send these channels only out the ports leading to the CobraNet® device to which they are addressed. Unlike multicast bundles, unicast bundles will not be transmitted unless a receiver is requesting that bundle. This allows destination controlled routing, where the receiver selects one of several possible transmitters to receive, and only the selected transmitter is activated.

It is possible to have far more than eight total bundles active on a switched network if most of those channels are sent unicast using unicast bundles. A given port on a fast Ethernet switch can only send eight bundles out without running out of bandwidth. Those bundles will consist of every multicast bundle on the network, plus any unicast bundle addressed to a CobraNet® device connected either directly or through other switches to this port on the switch.

Some switches have gigabit Ethernet ports in addition to the fast Ethernet ports. The gigabit ports can be used to transfer data between switches with 10 times the bandwidth of a fast Ethernet port and can carry ten times as many bundles as fast Ethernet can. Gigabit Ethernet also transfers data at ten times the speed of fast Ethernet, and thus can have as little as 1⁄10 the forwarding delay. This can become very important in larger networks.

Unlike repeater hub-based networks, CobraNet® over a switched network does allow coexistence with ordinary computer data on the same network, because there are no collisions with the audio. There is the possibility that CobraNet® traffic on the network will cause problems for 10 Mbit/s Network Interface Cards (NICs) used for computer data traffic. Recall that multicast bundles are sent to all switch ports in the same network. Since 8 bundles will fill a fast Ethernet (100 Mbit/s) switch port, if that port is connected to a 10 Mbit/s NIC (most fast Ethernet switch ports are dual speed 10/100 ports) then it is easy to see that multicast data from CobraNet® can saturate the 10 Mbit NIC and make it drop the computer data packets it needs.

There are several possible solutions: one easy solution is to upgrade the NIC to 100 Mbit/s full duplex.

1. Another possibility is to use little if any multicast bundles.
2. Most managed switches have multicast filtering features. These allow you to exclude multicast traffic from a specified port. If your data is carried by the Internet protocol (IP), it is usually safe to filter all multicast traffic except the FF:FF:FF:FF:FF:FF destination address used by the address resolution protocol (ARP) associated with IP.
3. Obviously separate physical networks for audio and data will solve the problem. Separate networks can also be created using VLANs, which are supported by most managed switches. All traffic in a given VLAN, even multicast traffic, is isolated to only those ports which are part of the VLAN. You can typically partition up to eight different VLANs, and assign ports to them as you wish. Uplink ports used to connect two switches can be connected to multiple VLANs, and the traffic from those VLANs is multiplexed onto that link, and then demultiplexed at the other end.

VLANs can also be used in some cases when you need to use more multicast bundles than is allowable on a given CobraNet® network. By splitting the network into two virtual networks you have the ability to run twice as many multicast bundles.

Another solution that can be used with some CobraNet® devices, is transmitting the same audio information on two, three, or four unicast bundles to specific destinations instead of a single multicast bundle. Please note that not all CobraNet® devices have this capability. Some devices can only transmit two bundles, while others can transmit four. Some devices only accept eight audio inputs, while others accept sixteen. Obviously if a device accepts sixteen audio inputs and can only transmit two bundles, it can’t use this technique.

Also be aware that different CobraNet® devices can receive different numbers of bundles, and select only certain audio channels from those bundles to use or output.

Follow this procedure when designing a CobraNet® network:

1. Make a list of all the audio sources and their locations.
2. For each source, list the destination(s) to which it needs to go.
3. Group the audio sources at a location into bundles with no more than eight audio channels in a given bundle (or seven if 24 bit).

4. Determine if each bundle can be unicast, or if it must be multicast.
5. Make sure you don't have more than four multicast bundles in a network. If you need more than four multicast bundles:

    - Consider using multiple switched networks or VLANs.
    - Consider transmitting several unicast bundles instead of one multicast bundle.
    - Use the following rules to see if you can send more than four multicast bundles on a given network or VLAN:
    - Carefully map the number of bundles sent to each port of the system. The total of multicast and unicast Bundles arriving at each switch port may not exceed eight.
    - If a half-duplex device that can only transmit two bundles, and is set to transmit using both its bundles is part of the network, then you must make sure that the network conductor is not transmitting a multicast bundle. This may require changing the default conductor priority of one or more devices in the system to assure this condition is met.
    - Map the bundles carried by every link in the system to make sure that the limit of 8 bundles each direction on a given fast Ethernet connection is not exceeded.

39.12.12 CobraCAD®

Fortunately there is an easier way to do steps 5 and 6 above. CobraCAD® can be downloaded for free from the Cirrus Logic Web site. CobraCAD® is a new software tool that provides a simple graphical user interface for the design and configuration of CobraNet® networks.

It allows you to draw your proposed CobraNet® network design using any of the CobraNet® devices on the market as of when the version you are using was released. You may also use any of a large selection of Ethernet switches.

After drawing the physical Ethernet interconnections, you next draw the bundle connections between the CobraNet® devices.

Then just press the Design Check button, and CobraCAD® will perform a design rule check. Designs that pass this check are extremely likely to work in the real world. There are still a few things CobraCAD® can't check for, so be sure to read the information and disclaimers in the Help system.

You may also want to check Cirrus Logic's CobraNet® Web site at: http://www.cobranet.info/ for the most recent version.

39.12.13 CobraNet® Hardware

At the heart of the CobraNet® interface, as shown in Fig. 39-57, is the digital signal processor, or DSP. It runs the software that together with the hardware provide all CobraNet® and Ethernet functions. It stores all the audio and Ethernet information as needed in stacks in the SRAM, converts the incoming synchronous audio into isochronous packets for transmission over the network, and converts isochronous packets from the network back into synchronous audio outputs. The DSP provides all interface functions to and from the device in which it is installed, and controls all other parts of the CobraNet® interface including the sample clock.

The sample clock is a voltage controlled crystal oscillator (VCXO), which is under the control of the DSP. If the CobraNet® interface is serving as the conductor, the sample clock is fixed in frequency and serves as the master clock for the network. In all other interfaces on the network, the sample clock is adjusted by the DSP so that it locks to the frequency of the network master clock.

The CobraNet® interface provides its clock signal to the device it is part of, but can also receive a clock signal from the device and use that signal as the network master clock if the interface is the conductor.

The CobraNet® interface can provide up to thirty two synchronous digital audio signals to the device, and accept up to thirty two synchronous digital audio signals from the device for transmission across the network.

The serial port can accept serial data which is then bridged across the network and appears at all other CobraNet® devices on the network.

The host interface allows bidirectional communication and control between the CobraNet® interface DSP and the processor of the host device in which the interface is located. Detailed information on the connections and signals on the CobraNet® interface to the host are available in the CobraNet® Technical Datasheet found in pdf form on the Cirrus Logic's CobraNet® Web site at: http://www.cobranet.info/.

39.13 Aviom

Aviom uses the Physical Layer of Ethernet. In other words it is transported over Cat5e cable with RJ-45 connectors. It does not use any other parts of Ethernet,

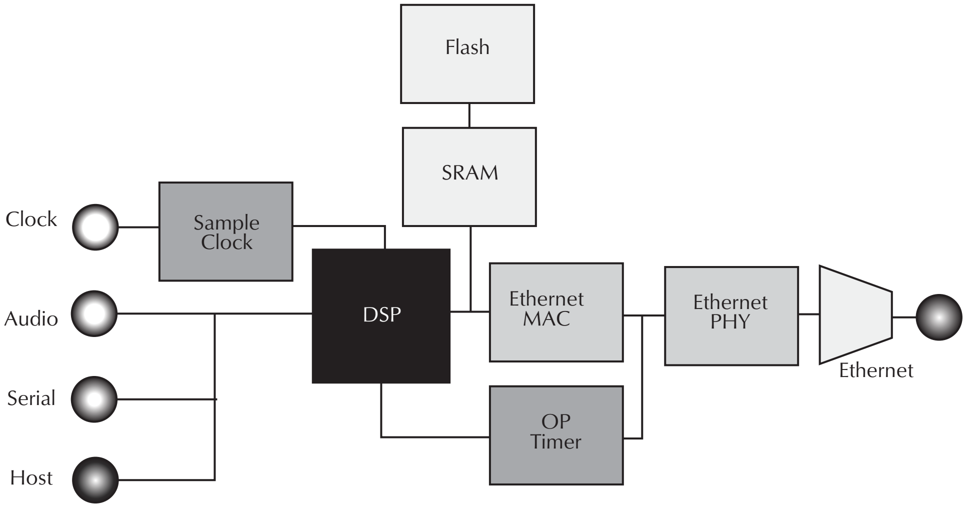


Figure 39-57. CobraNet® interface hardware. From the right there is the Ethernet network connection, and the isolation transformers. The PHY is a chip that acts as the Ethernet physical interface. The MAC is the Ethernet media access controller. The PHY and MAC chips together with the transformers constitute a standard fast Ethernet interface. The Flash memory serves as nonvolatile storage for the DSP firmware and the management variable settings. SRAM provides memory for the DSP processor. All the audio and Ethernet buffers are located here. The DSP (digital signal processor) is the heart of the CobraNet® interface and provides all control and processing functions. The sample clock is controlled by the DSP and either serves as the master clock or locks to the Conductor's master clock over the network. The OP timer controls transmission of packets onto the network.

but instead uses its own protocol. Aviom says on its Web site, "A-Net manages data differently than Ethernet does," which has advantages as well as disadvantages. Aviom at this time offers two different versions of its technology, Pro16 and Pro64. Pro16 is limited to point to point connections while Pro64 is more flexible. Pro64 allows up to sixty four audio channels and lets all devices see all sixty four channels. The Aviom protocols are low latency and simple allowing inexpensive but effective products such as its personal monitor mixers, which have revolutionized personal in-ear monitoring onstage and in studios. A single Cat5e cable to a small box by each musician carries sixteen audio channels and power, and allows the musician to make his own monitor mix exactly as he wishes.

39.14 EtherSound

EtherSound, as the name implies, does comply with the 802.3 Ethernet standard, but EtherSound networks are usually not built the same way Ethernet networks are. Ethernet networks are built using a star or star of stars topology, where each edge device connects to a switch port. Other than the simple case of a two device network, Ethernet edge devices do not directly connect to each other. Other than rare Ethernet edge devices with redundant ports, most Ethernet edge devices have only a single Ethernet port. Ethernet devices are never wired in a daisy-chain or cascade. EtherSound, on the other hand, provides in and out Ethernet ports on its edge devices, and in many cases builds networks by daisy-chaining its edge devices. This can result in simpler network designs, but also means that if a device fails in the middle of a daisy-chain it splits the rest of the devices into two isolated chains. EtherSound can also use switches in a more conventional star topology, but then devices downstream of the switch can't send audio back to the devices before the switch. The devices can be wired in a ring for fault-tolerance, and daisy-chain, star, and ring topologies can be mixed in the same network.

EtherSound has low latency and can support multiple sampling rates mixed in the same network. In order to get this low latency, EtherSound traffic must not be mixed with ordinary Ethernet traffic on the same network or VLAN. In EtherSound, 96 kHz streams occupy two EtherSound channels, while 192 kHz streams take four.

39.15 Dante

A new entry into the digital audio networking world is Dante from Audinate. Unlike the other real-time digital audio networking protocols, Dante makes use of the new IEEE 1588 real-time clocking standard to solve many of the issues facing those who would use Ethernet for audio transport. Dante also uses the standard UDP/IP data transport standards. This allows it to use standard Ethernet ports on a computer, for example, instead of requiring dedicated hardware to interface a computer to the audio network. Dante supports multiple latencies, sampling rates, and bit depths in the same network.

39.16 QSC

QSC Audio has introduced a new Ethernet-based digital audio networking scheme in some of its new products. It allows standard Ethernet ports on a computer to serve as audio transport ports, and does not require dedicated hardware for audio I/O. It is designed to take full advantage of Gigabit Ethernet and other advances in Ethernet technology, and to stay fully compatible with Ethernet as it evolves. Among the advantages it brings to audio networking are high channel counts, low latency, and the ability to operate over many switch hops. It uses automatic configuration techniques to greatly simplify the process of setting up an audio network and make it fast and easy.

References

1. AES-2id-2006—AES information document for digital audio engineering—Guidelines for the use of the AES3 interface.
2. AES3-2003—AES Recommended Practice for Digital Audio Engineering—Serial transmission format for two-channel linearly represented digital audio data.
3. AES-3id-2001 (r2006)—AES information document for Digital audio engineering—Transmission of AES3 formatted data by unbalanced coaxial cable.
4. AES5-2003—AES recommended practice for professional digital audio—Preferred sampling frequencies for applications employing pulse-code modulation.
5. AES10-2003—AES Recommended Practice for Digital Audio Engineering—Serial Multichannel Audio Digital Interface (MADI).
6. AES11-2003—AES Recommended Practice for digital audio engineering—Synchronization of digital audio equipment in studio operations.
7. AES18-1996 (r2002)—AES Recommended practice for digital audio engineering —Format for the user data channel of the AES digital audio interface.
8. AES42-2006—AES standard for acoustics—Digital interface for microphones.
9. IEC 60958-1 Ed3—Digital audio interface—Part 1 General.
10. IEC 60958-2 Ed3—Digital audio interface—Part 2 Serial copy management system.
11. IEC 60958-3 Ed3—Digital audio interface—Part 3 Consumer applications.
12. IEC 60958-4 Ed3—Digital audio interface—Part 4 Professional applications.
13. IEEE Std. 802.3, 2000 edition—Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications.
14. AES3-1992: The Revised Two-Channel Digital Audio Interface, by Robert A. Finger, *J. Audio Engi. Soc.*, Vol. 40, no. 3, March 1992.
15. Twisted Pair Cables for AES/EBU Digital Audio Signals, by D. G. Kirby, *J. Audio Engi. Soc.*, Vol. 43, no. 3, March 1995.
16. Controlling and Monitoring Audio Systems with Simple Network Management Protocol (SNMP), preprint 4760 (A-1), by Kevin P. Gross and Tom Holtzen, presented at the 105th AES Convention 1998 September 26–29, San Francisco, CA.
17. CobraNet® Technology Datasheet—11 April 2001—Version 1.1—Kevin Gross—Peak Audio.
18. Digital Audio Networking Demystified, by Brent Harshbarger, *Pro A/V Magazine*, 2007-08-01

This page intentionally left blank

Chapter 40

Message Repeaters

by Glen Ballou and Vic Cappetta

40.1 Digital Audio Storage 1515
40.2 Message Repeaters 1515
40.3 How Message Repeaters Work 1515
40.4 Message Repeater Usage 1516
  40.4.1 Hospital Applications for Message Repeaters. 1516
  40.4.2 Factory Floor Applications for Message Repeaters 1517
  40.4.3 Retail Store Applications for Message Repeaters 1517
  40.4.4 School Applications for Message Repeaters 1518
  40.4.5 Transportation Services Applications for Message Repeaters 1518
  40.4.6 Information/Broadcast Services Applications for Message Repeaters 1518
  40.4.7 Message-on--Hold Applications for Message Repeaters 1519
  40.4.8 Museum Applications for Message Repeaters 1520
    40.4.8.1 Inductive Loop Systems 1520
    40.4.8.2 Infrared Systems 1520
    40.4.8.3 RF Systems 1520
    40.4.8.4 Sophisticated Systems 1523
40.5 Narrow Beam Loudspeaker System 1524
  40.5.1 Voice Evacuation/Mass Notification Systems 1526
    40.5.1.1 Case Study 1526
40.6 Audio Archival 1527
Acknowledgments 1528
Additional Reading 1528

This page intentionally left blank

40.1 Digital Audio Storage

There are numerous applications for digital audio storage, that for day-to-day use, and that which is preserved for future generations. We will first address the applications and workings of message repeaters as they are used on a daily basis. Secondly, we will discuss the specifics of archiving audio information.

40.2 Message Repeaters

Message repeaters have come a long way in 40 years. The original repeater was a person sitting at a microphone making an announcement over a public address system at or about the right time. This system had both advantages and disadvantages. The message could be changed at any time, and if area switching was available, different messages could be sent to different areas in real time. Different messages could not be sent at the same time. Another disadvantage was requiring dedicated personnel 24 hours a day, seven days a week. In an emergency, an individual was required to stay at the microphone and announce in a calm and persuasive voice—a difficult thing to do at best.

With the introduction of tape recorders, messages could be prerecorded and played back manually or automatically. Unless a multichannel recorder or multiple recorders were used, only one message could be played at one time. To play different recorded messages required recording them in series and locating the desired message by fast forwarding the tape, time consuming in an emergency situation. With the design of lubricated tape, continuous loop tape recorders were used. This eliminated the requirement of rewinding but also meant the message could not be repeated until the tape had taken its course. Tape stretch and breakage was always a possibility. Auto-rewind cassette players were also used but had the same problem of reel-to-reel machines, they had to be rewound.

The late 1970s brought about the introduction of digital storage devices using solid state equipment. Digital message repeaters have the following advantages over the previous systems.

- Reliability.
- Flexibility.
- Solid state reproduction quality.
- User-recordable messages.
- Programmable, often user-programmable, locally or remotely.
- Remotely controllable.

40.3 How Message Repeaters Work

The use of digital signal processors (DSPs) has greatly simplified digital circuitry and made it possible to design digital message repeaters (see Chapter 39).

A digital message repeater in its simplest form is shown in Fig. 40-1. In this system, a permanent message is digitized by the manufacturer and stored in the digitized message storage circuit. To change the message requires sending the unit back to the manufacturer for reprogramming.

Upon contact closure, the control circuitry transfers the digitally stored message to the digital to analog converter (D/A) circuitry where it is changed into an analog circuit. The analog signal is only approximate at this time so a filter is used to smooth the waveform and limit the frequency response. Normally this type of unit has a frequency response from 300 Hz–3 kHz or the basic telephone response. The filtered output is then directed to the audio output circuitry where it is amplified, made balanced or unbalanced, and matched for the proper output impedance.

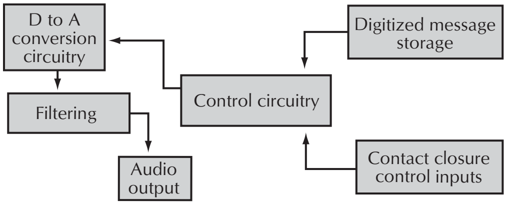

D to A conversion circuitry
Digitized message storage
Control circuitry
Filtering
Contact closure control inputs
Audio output

Figure 40-1. Digital message repeater.

An intelligent digital message repeater, such as an Instaplay™ by ALARMCO is shown in Fig. 40-2. In this system, messages can be recorded from a microphone (Mic), Aux. Input (AUX) (e.g., a CD or MP3 player), or standard touch tone telephone (control Phone). The analog input is then filtered and converted to a digital signal (A/D conversion circuitry) and stored (digitized message storage). Several thousand messages can be stored and individually replaced at any time. By using flash memory, sound quality is assured by storing audio data using 16 or 24 bit samples in the Instaplay™. Audio and programming data may also be downloaded digitally. With memory and intelligent firmware, each new recording can be longer or shorter than the original with all unused memory available for the recording.

To simplify recording from a control phone (either locally or remotely) [*Control Phone* and *Telephone Network Interface*], prerecorded instructions are stored in the announcer in digitized form [*Command Prompt Audio Storage*]. These instructions guide the user through each step. The telephone's keypad is used to respond to the prompts.

Installers can create a schedule for the playback of messages with the scheduler [*Internal Program Storage*] supplied by the manufacturer. Using this scheduler, the announcer can automatically switch from one playlist to another at various times throughout the day or week. The announcer insures not only that announcements are made, but also made on time.

With an intelligent announcer, the installer can specify [*User Program Storage*] where a message should be played [*Audio Outputs*] as well as the time delay between the announcements over each output channel.

Instaplay™ always knows when it's talking. By activating one of the numerous relays, [*Output Relays*], the installer is able to use the announcer to direct other activities, such as turning on lights or to trigger other announcers. The installer can control these relays by entries on a playlist [*User Program Storage*].

During playback, selected messages can be played in any order (including repeats), as specified by the installer [*User Program Storage*]. Background music, such as Muzak, can be played between messages [*Music Feedthrough*] with many message repeaters. An intelligent message repeater, however, can interrogate the software [*User Program Storage*] and decide whether to duck down or mute the background music during each individual message.

With intelligent message repeating messages can be triggered externally in numerous ways [*Contact Closure Control Inputs, Serial Communications Link, Control Phone*], and to play a message sequence or queue new messages. In addition, intelligent repeaters allow customers to record new messages, modify schedules, playlists and other programming parameters either locally or remotely. When accessing the announcer remotely, [*Telephone Network Interface*], a security code can be employed. A queue is a sequence of message files that has been selected to be played through a particular channel. A playlist is a command list that queues the messages to play in a defined sequence and channel and includes the ability to operate external devices. It also has the ability to start and stop messages from an external trigger.

Instaplay™ gets its native intelligence from its embedded firmware [*Internal Program Storage*]. Supplied by the manufacturer, these programs not only define the internal operation of the machine, but also define default parameters, such as how often to repeat a recorded message or how to act when a request is received.

Instaplay™ gets its application intelligence from the programming entered at the job site and stored in the announcer's random access memory [*User Program Storage*]. The installer must easily be able to change these default parameters within the announcer to have it operate as the application dictates. Whereas a casino operation may want to announce hundreds of events throughout a particular day according to a predetermined schedule, the announcement of a train's arrival at the platform needs to correspond with the actual arrival time, rather than the scheduled arrival time.

The microprocessor must always coordinate both the default system values and the user-specified parameters to operate appropriately for different applications.

40.4 Message Repeater Usage

Message repeaters are used in many venues:

- Hospitals.
- Schools.
- Factories.
- Amusement Parks.
- Retail Stores.
- Tourist Attractions.
- Transportation Services.
- Information Providers/Broadcast Services.
- Message on Hold.
- Museums.

40.4.1 Hospital Applications for Message Repeaters

Message repeaters have been sold into numerous hospitals. One common application is to broadcast different messages into separate locations in the hospital. The messages may be divided into all area announcements, or those that are announced only in public areas or patient areas. For example, visiting hour reminders are broadcast into all areas of the hospital, including patient's rooms. No smoking reminders might be announced only in lobbies, cafeterias, and waiting rooms, while doctor calls would only be announced in patient areas.

In addition, with a scheduler, message repeaters can be used to announce when visiting hours are about to end and when visiting hours begin again.

Hospitals around the country are trying unique ways of increasing their patient satisfaction. Many hospitals are now using message repeaters to play "Brahms Lullaby" when a baby is born. It adds a little smile to everyone's face.

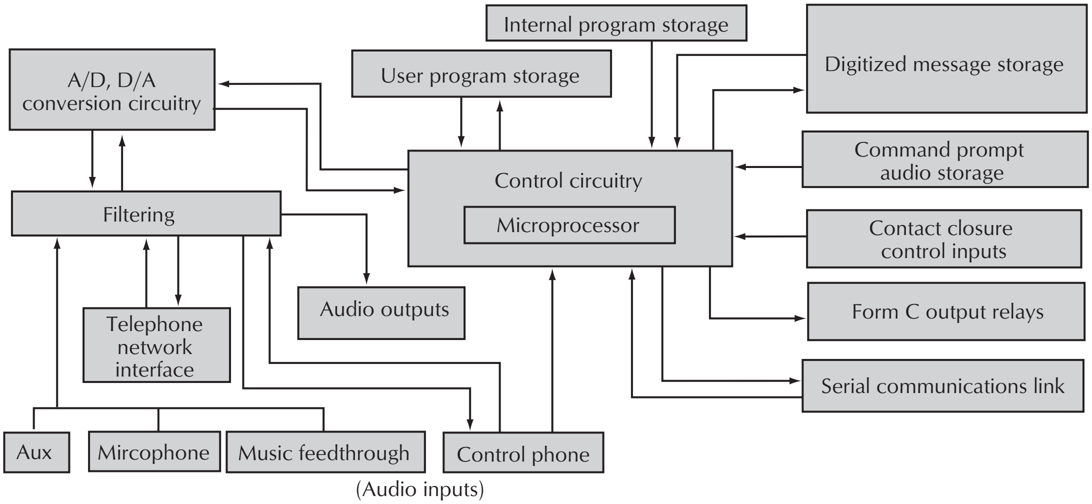

Internal program storage
A/D, D/A conversion circuitry
User program storage
Digitized message storage
Command prompt audio storage
Control circuitry
Microprocessor
Filtering
Contact closure control inputs
Audio outputs
Form C output relays
Telephone network interface
Serial communications link
Aux
Mircophone
Music feedthrough
Control phone
(Audio inputs)

Figure 40-2. An intelligent message repeater. Courtesy ALARMCO.

A message repeater can also be connected to controlled doors such as in geriatric and psychiatric wards to tell nurses of unauthorized door openings. Voice messages can announce immediately throughout the building which door is open, so personnel do not have to go to a central alarm panel to see which door is ajar, a much more effective and faster method than visual indications.

When a Code Blue Dispatch happens, message repeaters can instantaneously repeat the message as often as required without requiring dedicated personnel.

40.4.2 Factory Floor Applications for Message Repeaters

Just-in-time manufacturing is a very popular way for companies to cut expenses. Some companies have documented savings of over one million dollars annually in reduced inventory expenses by having components delivered where required just-in-time. This way, unused parts don't require valuable and expensive floor space, nor does the production line ever slow due to lack of components. Many of the largest factories in the country, such as Motorola, General Motors, and Xerox, have been practicing this method of cost reduction for many years.

Message repeaters can be used to broadcast a message when certain parts need to be restocked. Messages can be triggered manually by having the assembly operator push a button, or automatically with a sensor to trip the message when the weight of the bin containing components becomes too light, sending a verbal message to stock.

Repeaters can also be used to announce lunch time, scheduled breaks, safety reminders, company announcements, and so forth. They can be configured to know when someone is entering (versus exiting) a hard hat area. A warning message can be announced on entering. A message reminding visitors to return their hard hats and protective eyewear can also be announced on exiting the area.

Another interesting use of a message repeater is the elimination of acoustical feedback in noisy environments. By recording the message or page on a message repeater and playing it back as soon as the recording is finished, the microphone-amplifier-loudspeaker-room feedback loop is broken and feedback is eliminated and pages can be automatically repeated. With PageDelay™ from ALARMCO, pages can be monitored as they are being recorded. Inappropriate pages can easily be canceled before they are aired.

40.4.3 Retail Store Applications for Message Repeaters

Message repeaters are ideal for in-store assistance requests. In-store assistance is becoming increasingly popular with retailers because it allows them to cut the expense of large staffs while still delivering service to customers who need it. Often the signs asking customers to "Press the button for assistance in this department" is connected to a message repeater.

Convenience stores have a captive audience when it comes to customers who are pumping gas. Message repeaters can trigger different advertising messages when someone drives up to the gas pump. Studies have proven that sales increase dramatically with this type of advertising. Repeaters can be configured to play the appropriate message to shoppers as they enter or exit a store.

Message repeaters are used in some stores to do targeted in-store advertising. Special repeaters allow messages to influence the customers when they are near the point-of-purchase. Targeting is done by using sensors to trip the control circuit, thereby giving advertising messages when someone enters a particular aisle or department. Even more specific targeting can be accomplished, for example, when someone reaches for a brand name product, as detected by a motion sensor, a message can be played to try the store brand.

Customized modes of operation form an elegant solution for in-store use. Specific options could include individual timers that would allow each department manager to specify the amount of time before the message repeater transmits another message triggered from the same department. A second audio output channel allows office or security personnel to know when someone needs assistance in sensitive areas, such as someone in the area carrying a handgun. Message repeaters can allow any department zone to be enabled or disabled from a central location, such as the front office. If a child is playing with a button, the store manager can temporarily disable that button until the child moves on.

With a built-in scheduler, message repeaters can announce appropriate closing times each day. Sophisticated software allows a single intelligent message repeater to perform many of these functions simultaneously.

40.4.4 School Applications for Message Repeaters

School closing or shortened day announcements for snow days etc. can be recorded and played back over the school PA system and the local town access TV channel.

Also class change announcements and activities announcements can be prerecorded and transmitted day or night automatically through the internal scheduler. A message repeater can also be set up as a dial-in line for sports schedules, latest news, and so forth.

Often school gyms or auditoriums are used after hours. Due to fire regulations, many do not have security gates to separate the used area from the rest of the school. A message repeater, triggered by a light beam or motion detector, can energize a message repeater to notify people that they are walking into a closed area, while simultaneously alerting personnel or security staff. In the event of a lock down situation, message repeaters quickly and reliably direct students and staff in safety measures.

40.4.5 Transportation Services Applications for Message Repeaters

To ease the burden on the driver, bus route stops can be manually selected, controlled, and played on a message repeater or they can be completely computer controlled. Message repeaters on tour buses are often used to eliminate the necessity of a tour director and to give a running dialogue of the tour route. Usually the driver controls the repeater by pushing a switch when he or she is ready for the next message to be announced.

Message repeaters on mass transit loading platforms can announce train arrivals and departures, safety messages, and upcoming schedule changes. The Statue of Liberty Ferry chose *Alarmco*'s Instaplay™ to intelligibly announce tour information and required safety messages, Fig. 40-3.

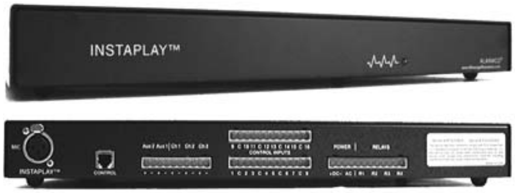


Figure 40-3. Instaplay message repeater. Courtesy ALARMCO.

40.4.6 Information/Broadcast Services Applications for Message Repeaters

USIA Voice of America uses message repeaters in a number of European countries to retrieve messages that are broadcast over a worldwide satellite network. The broadcast messages are downloaded in various languages into the appropriate repeater. Local radio stations in the individual countries dial into the repeater and download the messages, which are in turn broadcast over the radio stations. Intelligent message repeaters have the ability to record remotely from a line level input, thereby allowing the remote telephone to control them, while they are recording from the satellite.

Because messages can be easily changed, they are ideal for radio broadcast directions, Traveler Advisory Radio (such as AlertAM from Information Station specialists), news broadcast services, and visitors information services. They are also widely used for employment hotlines, movie theater schedules, sports score lines, and weather information lines.

40.4.7 Message-on--Hold Applications for Message Repeaters

Telephone on-hold equipment is operating whenever we hear anything but silence when we are put onhold. There are three basic categories of on-hold equipment:

- Tuners/radios, where the on-hold program is a radio broadcast.
- Cassette players, where the on-hold program is on an endless loop cassette.
- Digital playback systems, where the on-hold program is stored on solid state memory chips.

Each system has advantages and disadvantages. While the tuner/radio is probably the least expensive method of playing on-hold programs, employing radio programs on-hold has a few pitfalls. For instance, radio programs include licensed music and retransmission is illegal for all but very small businesses without paying licensing fees. The tuner/radio might also be transmitting competitive ads.

Cassette players allow end users to either make their own on-hold program, or have it made by a professional on-hold studio. The disadvantages to cassette players are tape and head wear from dragging the tape across playback heads and the necessity for head cleaning and demagnetization every few weeks to slow head wear.

Digital playback units came into use in the 1980s, primarily because of their high reliability. When the program is loaded into solid state digital memory chips, it can be played back continuously with no moving parts and no wear and tear. The program sounds the same on the millionth play as it does on the first.

The programming and messages for on-hold players can be recorded either locally or remotely. Local recording and programming requires someone on-site to record the message or load cassette tapes. Some of the more intelligent systems allow for remote downloading of the programming and messages using satellite systems, FM subcarrier audio channels, a standard telephone line, a modem, or the Internet. During playback, these remote download units have all the reliability advantages of other digital on-hold equipment. They have the additional advantage over conventional tape download equipment that they are completely hands-off at the installation site. They require no intervention from on-site store personnel, who may be unwilling or unable to load in the tape.

Digital on-hold players employ memory chips to store the program. To produce a frequency response from 20 Hz–20 kHz would require a tremendous amount of memory. Telephone lines normally will not transmit a greater frequency response than 300 Hz–3.5 kHz, therefore it is not practical to increase the response and the equipment cost to cover a much wider range.

As is the case with CDs, digital on-hold units sample the incoming signal many times a second to store the signal into digital memory. Theoretically, the greater the samples per second—i.e., the sampling rate—the better the sound quality. Sampling rate is usually expressed in kilobits per second (Kbps). Toll quality telephone performance (the best performance any telephone network will allow) is 64 Kbps so there is no need to produce on-hold units with a sampling rate greater than 64 Kbps.

Sampling rate is only one measure of the audio quality of a digital downloadable on-hold unit; a network of filters and frequency compensators also contributes to the sound quality.

In most cases, nontechnical employees connect the on-hold equipment to the line and load new taped music/messages. Some units require no controls, no level setting, and no start/stop control because the units employ microprocessors to control all aspects of the download/play process.

The Bogen HSR series unit is an example of a full microprocessor-controlled on-hold system. Various models have a 4, 6, 8, or 12 minutes of memory capacity. The HSR's automatic operation assesses the start and stop point of the audio, sets record levels, downloads, and goes into play mode automatically. The unit also incorporates a one-play trigger mode for making a single message such as store closing.

The Mackenzie Laboratories, Inc. Dynavox series are on-hold systems that can also can be used as storecasters. One series has a 3.4 kHz bandwidth for telephones and the other series has a 6.8 kHz bandwidth for storecast and other wide-band requirements. The bit rate increases from 96 to 196 Kbps and the sampling frequency increases from 8 to 16 kHz with the increasing frequency response. Audio storage requires 16 MB DRAM (dynamic random-access memory) to record 32 minutes at 96 Kbps and 16 minutes at 196 Kbps. These units have a noise floor and dynamic range of greater than 70 dB.

Intelligent message repeaters, such as the Instaplay™ series by ALARMCO, can provide on-hold music and messages, storecasting messages, triggered customer assistance messages, and automatic store closing announcements—all with a single announcer. In addition, this type of repeater can also play existing background music through, thereby eliminating the need for customers to purchase a cassette with recorded music.

40.4.8 Museum Applications for Message Repeaters

Message repeaters with multiple inputs are ideal for museum displays that require several messages associated with a single display. For example, messages that are tailored to either adults or children, short messages or detailed messages, or multiple language messages.

Museums can be exciting with interactive displays and audiovisuals replacing the static placards of the past. Today visitors are accustomed to a multimedia environment where programming captures their attention with entertainment mixed in with an educational message.

The original museum audio systems included narrations and sound effects. Loudspeakers were mounted in front of the exhibit, often at knee level. The source was often continuous loop tapes that could not be rewound and they ran continuously so if the listener came in during the message he or she would have to listen to the end before the beginning. If the tape had automatic stop at the end of the narration, the listener could push the start button and hear the message from the beginning. Of course anyone coming in after the dialogue was started would have to listen to the end before restarting it. Often multitrack playback tape machines were used so the listeners could pick their language of choice. While not the best system, it was a step beyond placards.

Around 1957, visitors carried a reel-to-reel tape recorder over their shoulder. With this system they not only heard the message but also got their exercise for the day.

40.4.8.1 Inductive Loop Systems

Another early system, and still used by some museums today, is to transmit the signal on a wire inductive loop antenna that surrounds the audience area. The listeners wear a receiver and earpiece and as long as they are within the boundary created by the loop, they can hear. As they step outside of the loop, the signal disappears. They can then go to the next exhibit, step into its loop, and hear the dialogue. An advantage of the system is it is simple and reliable and works with hearing aids. The drawbacks are:

- Poor frequency response making it useful only for voice.
- As the signal is analog and operates much like an AM radio station, the volume and sensitivity vary with the distance of the listener to the loop.
- Affected by external electrical noise such as lightning, electric motors, and SCR lamp dimming circuits.
- Requires a wire loop around the area of interest, sometimes rather difficult to install and hide.

For more information on magnetic induction loop systems, see Chapters 41.2 and 42.2.1

40.4.8.2 Infrared Systems

Another type of system by Sennheiser and others uses infrared (IR) transmission. In this system the message is transmitted via wireless infrared using amplitude and frequency modulation processes. They come in either narrow band for multichannel setups or wideband for high quality.

The area of reception is confined to line-of-site or an individual room. Through reflections, however, it can bounce around corners into other unwanted areas. While they can cover large areas effectively, they are limited when it comes to multiple exhibits in a confined area. Another problem with IR is its poor operation in the sun or very bright areas.

Dual channel systems normally operate subcarriers of 95/250 kHz or 2.3/2.8 MHz. The emitters are placed around the room to give even coverage and they may be daisy-chained for easy installation.

For more information on infrared systems, see Chapters 41.3 and 42.2.5.

40.4.8.3 RF Systems

Today RF systems are most often used, making the systems much more versatile and simpler to install. These systems range from simple to quite complex. The following systems are only a smattering of what is available but give an indication of the features available.

Acoustiguide has been in business for over 50 years. Its major system is the Acoustiguide 2000 Series and includes three AG 2000 players—the *Wand*, the *Mini*, and the *Maxim*.

All three systems use MP3 and Windows Media Audio 4.0 for full production sound, and their own software called Vocoder for voice only.

The *Wand* is 12.5 inches long, 2.5 inches wide, and 0.75 inch deep and weighs 9.2 oz, Fig. 40-4. It can hold up to 500 selectable languages or programs or up to 8000 messages. The controls include Play, Clear, Pause, Fast Forward, Rewind, Volume Up, and Volume Down. It can play for 12 h continuous without charging and can accommodate surveys, games, and educational question-and-answer formats. Battery charging can be accomplished in 3 hours in the charging/programming rack.

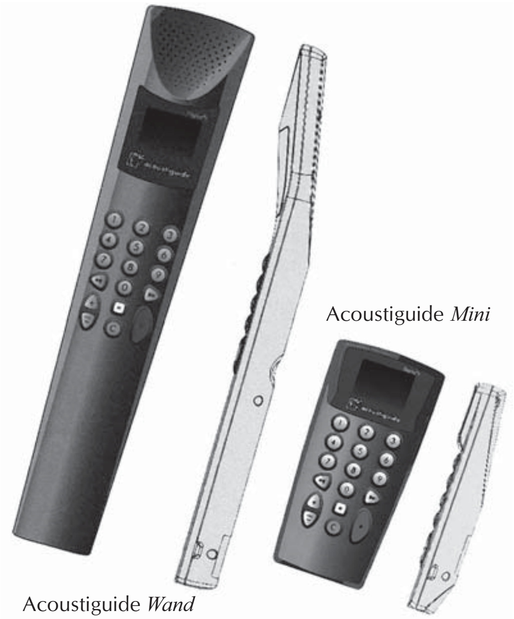


Figure 40-4. Typical wands used in museum systems. Courtesy Acoustiguide Inc.

Because of the design of the *Wand*, it is easy to encourage corporate sponsorship including rotating logos on the LCD screen and the flat areas on the casing are good for applying logos and graphics.

The *Mini* has many of the same features as the *Wand*. The *Mini* is ideal for highly produced audio programs that blend narration, archival audio, large interviews, music, and sound effects making exhibits come to life. The *Mini* comes with headsets or single earpieces. The unit is 5.6 inches long, 2.6 inches wide and 0.75 inch deep and weighs 5 oz. The controls are the same as with the *Wand*. It will also play for 12 h without charging and can be fully recharged in 3 h.

The *Maxim* can hold 200 h of stereo sound or over 2000 h of voice in either linear, random access, or combination tours. It can hold 500 different programs and over 12,000 messages on each unit to provide tours on different subjects or foreign languages. The unit is 7 inches long, 3.9 inches wide, and 1.5 inches deep and weighs 15 oz. It has the same controls as the *Wand* and the *Mini*.

The Acoustiguide storage racks recharge the batteries and include a programming card that is about the size of a credit card. The programs can be either written by the client or by Acoustiguide which can provide creative and production services. The programs are downloaded from the Internet or from CDs onto a laptop computer. As new material is written, recorded, and digitized, it is put on the program card, which automatically updates the players as they are being charged.

To operate the system the visitor is given a player. A staff member sets up the player for the language and the complexity of the tour. The tour could be long or abbreviated to control traffic when the museum is crowded or can be set up for adults or children. The visitor can adjust volume at each area to compensate for noise level. When the visitor is at an exhibit, he or she punches in the number corresponding to the exhibit as shown on a placard. The visitor can then pause the program, rewind it, or fast forward it.

Acoustiguide's newest unit is a compact screen-based player, developed and designed specifically for on-site interpretation of museums and visitor venues. The Opus series allows institutions to provide visitors access to various digital resources—video, images, and animation, plus the traditional audio, Fig. 40-5.

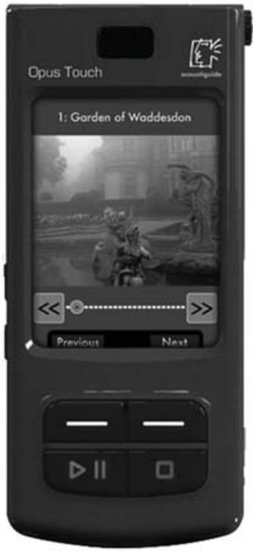


Figure 40-5. Acoustiguide Opus Touch™ screen-based player. Courtesy Acoustiguide Inc.

The high-performance computing capabilities of Opus include sophisticated graphic images and digital movies as its processing speed and memory capacity enable delivery of high-resolution video files and CD-quality sound.

The administrative user interface allows for simple additions and deletions of content, as well as more complex functions, such as integration of timed audio with video and images.

Opus comes in two formats: Opus Click™ and Opus Touch™. The click format utilizes a keypad while the touch format utilizes a touchscreen. The systems incorporate the following:

- Remote triggering and synchronization.
- Remote content activation via IR and RF technologies.
- Synchronization to external multimedia or show control systems.
- Data collection and visitor surveys.
- Software tracks user click stream.
- Customized surveys can be incorporated into the guide's audio/visual content.
- Produces easy-to-read reports.
- Visitors can bookmark items of interest either for on-demand printing via MyCollection™ or other postvisit services such as e-mailing information home.
- Map-driven or object-driven modes.
- Full color TFT LCD screen.
- Large, expandable memory.
- Compatible with a complete range of audio, video, image and animation multimedia formats.
- Dual-listening mode, via internal speaker and/or through headset/earpiece.
- Up to 12 h of usage between charges.
- Remote activation via IR and RF.
- Opus Content Management System—easy setup and install, can be used by client.
- Dual-listening mode.
- A fold-out loudspeaker, each player can be used either as a true wand or as a headset unit.
- Headphone integrated into strap.
- MP3 stereo sound quality.
- Range of sixteen stepped volume levels.
- 500 h of multilingual audio content.
- MP4 and JPEG visual quality.
- QVGA resolution (320 × 240 pixels) and 65,536 color depth.
- 28 h of video, or 10,000 images.
- 2 GB memory; expandable.
- Holds multiple languages and tours.
- Guides can contain any combination of audio, images, animation, or video clips.
- Graphically rich user interface with menu selection and navigation functions.

Another system is made by Tour-Mate. Its SC500 Listening Wand is 13 inches long, 1.8 inches wide, and 1 inch deep and weighs 8 oz. A carrying strap is attached inside the wand for added strength and so it that it cannot be unclipped by the user. The system is powered by a rechargeable nickel metal hydride battery pack which will deliver 10 h of continuous play from a full charge and can be recharged in 3–4 h, Fig. 40-6.

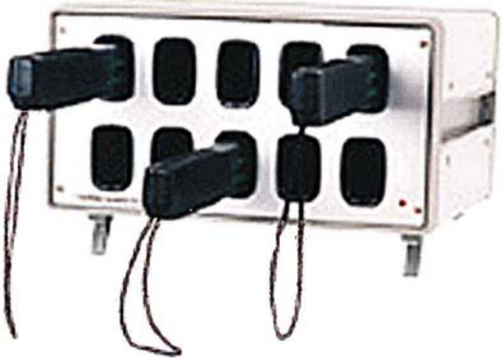


Figure 40-6. Typical charger/programmer.

The maximum capacity of the system is 24 h of mono sound or 12 h of stereo sound. The message can be expanded on-site. The wand can store several tours and/or versions of the tour. The software permits a staff member to type in a code that locks out all tours but the desired one. A keystroke permits one to see what version of the tour has been selected.

The *Tour-Mate editing capability* software is windows compatible. The editing software permits the user to input tour messages or message segments and to perform such functions as: cut, paste, parametric equalization, normalization, variable gain, variable compression, insert message queues, and program message sequences.

MyGuide by Espro is a system much like the previous two. It uses a wand that has the tour narration downloaded from the storage-rack/power supply recorded through a flash memory card. This system runs for 10 h between charging and can have up to 4 h of audio capacity. The bandwidth is between 300 Hz and 4 kHz so it is particularly useful for voice only.

ExSite MP3 system by Espro can have up to 72 h of multilingual content, uses a wide alphanumeric and graphic LCD screen, and can be synchronized to external multimedia presentations such as DVD and video. It can also collect and analyze visitor usage data.

The system can be used with the unit's built-in speaker or with plug-in earphones.

GroupGuide by Espro is a portable system for group tours where the visitor wears a personal receiver with headphones, and the guide wears a transmitter with microphone.

AKG, Sennheiser, and Williams Sound also have tour systems using wireless microphones and wireless receivers. All of these systems as the GroupGuide are useful for guided tours.

40.4.8.4 Sophisticated Systems

One sophisticated system is *GuidePORT™* by Sennheiser. To operate GuidePORT, the museum is set up into zones or cells, Fig. 40-7. These zones may be separate rooms or floors or a section of a large room. Audio files associated with the exhibit in a cell and their corresponding identifier unit are created and/or stored on a standard PC. The files are uploaded by GuidePORT software to multichannel RF (radio frequency) wireless cell transmitters located in each individual cell. Each cell transmitter stores the audio for its particular zone. The audio (prerecorded and/or live stream) for that particular cell is downloaded into the visitors' receivers when they enter the cell.

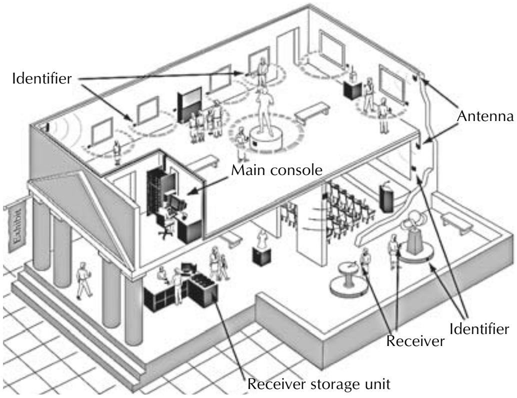


Figure 40-7. GuidePORT system layout for a typical museum. Courtesy Sennheiser Electronics.

GuidePORT's charger system can store and charge ten wireless receivers. Chargers can be linked to accommodate 5,000+ receivers. Receivers can operate for up to 8 h between charges. The charger system is linked to the control unit (PC) to allow programming the receivers for language and/or level.

To enable management of a frequently changing exhibit environment, Sennheiser has engineered a list-based audio configuration software so the museum management can control the audio tours by simply updating the master audio list as exhibit items and corresponding identifiers are moved.

Discreet wireless antennas are strategically placed throughout the exhibit to allow receivers and cell transmitters to interoperate. The system is designed to operate on license-free dedicated radio frequencies in the 2.4 GHz ISM band that are ideal for digital audio and resistant to outside radio interference.

Battery-operated or externally powered wireless identifier units are hidden near or behind each exhibit. The wireless architecture behind the GuidePORT system allows for quick and easy setup of the museum because, as an exhibit is moved, the associated identifier is moved along with it making it easy to rearrange an exhibit space.

The visitors are given a lightweight receiver that fits in the palm of their hand or can be hung around their neck, and a headset. The receiver is programmed by a staff member to the language and level the visitor desires. The system is hands free so the visitor is not required to press buttons to match exhibit placards. The visitors proceed into the exhibit at their own pace and as they move from exhibit to exhibit, the system automatically dissolves the audio from the previous message to the new one. Visitors can adjust the volume and pause or repeat information they would like to hear again, Fig. 40-8. The headphones fit all age groups and can include a sponsor logo.

When the visitor enters a zone, audio files for all of the exhibits within that zone are downloaded into the visitor's receiver. The identifiers automatically trigger the receiver when a visitor is within a specified range of the displayed item to play the corresponding audio file. The trigger range, along with other parameters of the identifier, can be programmed via an infrared enabled Palm™ compatible PDA.

GuidePORT can integrate live audio into the presentation. The visitor can listen to live demonstrations, concerts, movies, and video presentations with synchronized sound just by walking into the area. The visitor can leave the area and walk to a new area and the audio program will automatically change.

All stationary components of GuidePORT are located in a central location. Cell transmitters interface with their base station PC via USB ports. A larger facility can network multiple base station PCs, including through an existing network. Antennas are connected using standard shielded Cat5 cable. Audio files may be created anywhere in any standard formt, which are then converted to .WAV files before they are

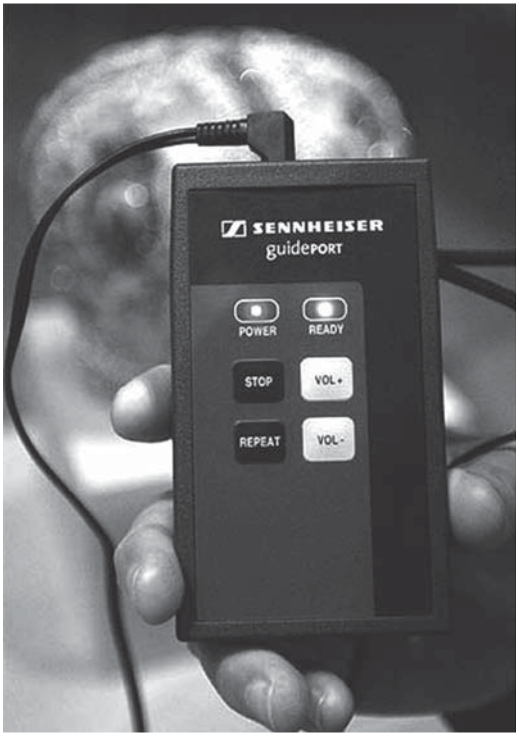


Figure 40-8. GuidePORT receiver. Courtesy Sennheiser Electronics.

imported into the GuidePORT system. The base station PC and/or central control unit is only needed when configuring or reconfiguring the system, and could be substituted with a temporary PC or notebook.

40.5 Narrow Beam Loudspeaker System

The directivity (narrowness) of any wave producing source depends on the size of the source compared to the wavelengths it generates.* Audible sound has wavelengths ranging from a few inches to several feet, and because these wavelengths are comparable to the size of most loudspeakers, low- to medium-frequency sound (20 Hz to 10 kHz) generally propagates omnidirectionally. Only by creating a sound source much larger than the wavelengths it's producing can a narrow beam be created. To accomplish this with standard loudspeakers would require loudspeakers 50 ft in diameter. A narrow beam of sound from a small acoustic source is accomplished by generating a beam of ultrasound, which becomes audible as it travels.

Ultrasound, whose wavelengths are only a few millimeters long, are much smaller than the source, and consequently travel in an extremely narrow beam.

Ultrasound contains frequencies far outside of our range of hearing, and is completely inaudible, but as the ultrasonic beam travels through the air, the inherent properties of the air cause the ultrasound to distort (change shape) in a predictable way. This distortion gives rise to frequency components in the audible bandwidth, which can be accurately predicted, and therefore precisely controlled. By generating the correct ultrasonic signal, we can create, within the air itself, essentially any sound desired.

Note that the source of sound is not the physical device you see, but the invisible beam of ultrasound, which can be many meters long. This new sound source, while invisible, is very large compared to the audio wavelengths it's generating, so the resulting audio is extremely directional, just like a beam of light.

Often incorrectly attributed to so-called Tartini tones, the technique of using high-frequency waves to generate low-frequency signals was in fact pioneered by physicists and mathematicians developing techniques for underwater sonar over 40 years ago.

Dr. F. Joseph Pompei, then a researcher at MIT, solved the problems of using ultrasound as an audible source that plagued earlier researchers. His design of the Audio Spotlight®* sound system has become the very first, and still the only, directional loudspeaker system which generates low-distortion, high-quality sound in a reliable, professional package, Fig. 40-9. Fig. 40-10 shows the sound field distribution with equal-loudness contours for a standard 1 kHz tone. The center area is loudest at 100% amplitude, while the sound level just outside the illustrated beam area is less than 10%.

Audio Spotlight systems are much less sensitive to listener distance than traditional loudspeakers, but maximum performance is attained at roughly 1–2 m (3–6 ft) from the loudspeaker.

Typical levels are 80 dB SPL at 1 kHz for the AS-16, and 85 dB SPL for the AS-24 models. The larger AS-24 can output about twice the power and has twice the low-frequency range of the AS-16.

The most common use of the Audio Spotlight system is to deliver sound to a specific, isolated area. Just as with lighting, the Audio Spotlight system is best

---

* Much of this section was copied with permission from copywrited text by Holosonic Research Labs, Inc.

* Audio Spotlight is a registered trademark of Holosonic Research Labs, Inc.

Figure 40-9. Audio Spotlight AS-16B system. Courtesy Holosonic Research Labs, Inc.

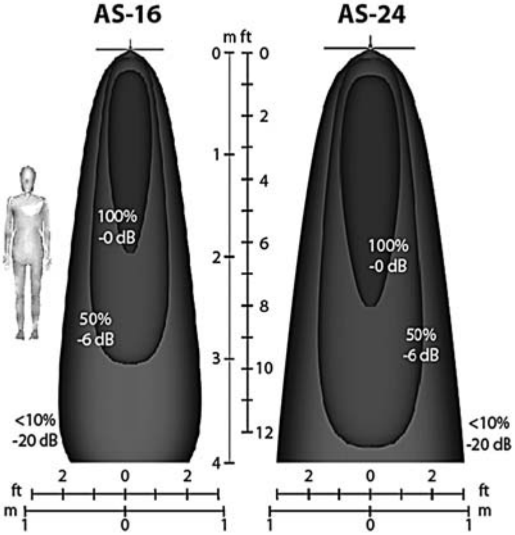


Figure 40-10. Sound field distribution of the Audio Spotlight AS-16 and AS-24 systems. Courtesy Holosonic Research Labs, Inc.

mounted directly above the listener, aimed downward Fig. 40-11. which provides maximum localization. The speaker panel can also be mounted on a wall, and angled downward, to reach the listener.

Multiple Audio Spotlight systems can be used to create a larger field of sound, or to increase the sound intensity in a given region. Just like visual spotlights, beams of sound can be aimed next to each other, to shape the sound field Fig. 40-12A, or multiple speaker panels can be aimed to one position Fig. 40-12B. Just as with light, sound from these systems will combine to increase output substantially.

While the beam generated by the Audio Spotlight system is very narrow, the beam will reflect from surfaces (and listeners) in your environment. To sound waves, solid surfaces are much like mirrors are to light. Therefore, to reduce reflections, an acoustically absorbing surface (such as carpet, padding, or curtains) should be used to catch the beam and reduce the reflection. Generally, this is most important only in very quiet spaces, where there is little background noise to mask minor scattered energy. Also, like light, reflections can be used as projection of audible sound. By directing the beam against a surface, one can create very interesting virtual loudspeaker effects. The beam will generally maintain its directivity after projection, so it is best to insure that the listener is in the path of the reflected beam.

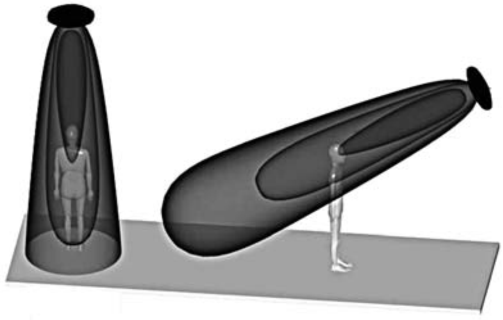
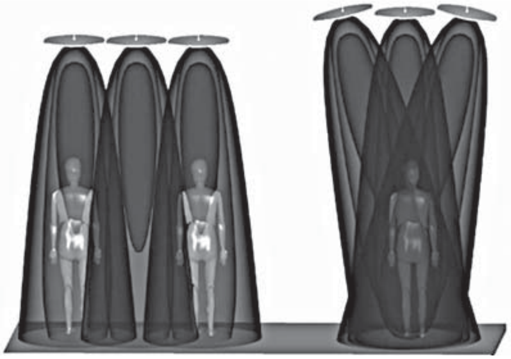


Figure 40-11. Mounting angles for Audio Spotlight systems. Courtesy Holosonic Research Labs, Inc.



A. Shaping coverage. B. Increasing SPL.

Figure 40-12. Multiple Audio Spotlight systems. Courtesy of Holosonic Research Labs, Inc.

The loudest sound area is directly in front of the speaker panel at a distance of 1–2 m. Reasonable listening areas are within the darker zones. Sound levels outside the beam are down by over 90%. In all sound systems, audibility is determined by sound level received versus background noise levels. Therefore, the beam will be perceived as more narrow in the presence of background noise, as any scatter from a listener or floor will be inaudible. This is much like the difference in shining a flashlight in a completely dark room versus one with background lighting.

40.5.1 Voice Evacuation/Mass Notification Systems

by Vic Cappetta, Cooper Notification

A logical evolution for message repeaters is the development of voice evacuation systems for life safety applications.

Cooper Notification, with the company's brands Wheelock®, Safepath® Waves®, and Roam Secure®, offers complete solutions consisting of supervised notification and audio systems, RF control, and network alerting.

Depending on the application, Voice Evacuation and Mass Notification are terms that are often used interchangeably. The term *voice evacuation* is traditionally used in the fire alarm industry; the term *Mass Notification* is more recent and has its origins in military installations and an ever-growing presence on college campuses.

The term Mass Notification has been adopted by the U.S. Army Corps of Engineers and is specified and defined in a driving document known as the UFC (Unified Facilities Criteria). Mass Notification requirements are detailed and specific, including intelligibility performance. The intelligibility aspect is gaining more and more momentum as a way of defining clear communications over a loudspeaker system. Basically, intelligibility measurements deal with a percentage of loss of consonant definition (%*Alcons* method) or may be measured on CIS (combined intelligibility scale) or STI (speech transmission index) platforms.

Voice evacuation is self-explanatory—a system broadcasts recorded or live voice announcements over a loudspeaker system, typically within a building. Mass Notification, which can be the same thing, can also extend to outdoor areas; for example, military bases or university campus environments. The differences can be subtle: the term Mass Notification is more often used when the system is used for more than simple fire messages; for example, severe weather, industrial incidents such as noxious gas release, or bomb threats.

If a Voice Evacuation system is specified as a Fire Alarm system, the system must be UL listed for fire, and the entire system must be monitored for integrity (supervised).

This means that all internal aspects such as power supplies, voice modules, etc., as well as external aspects such as speaker loops must be monitored for shorts, opens, or ground fault. These conditions must be reported as troubles to the system headend.

The UFC basically defers to NFPA (National Fire Protection Association) for its technical requirements; therefore Mass Notification systems are typically supervised and resemble fire alarm notification systems in many respects. Some differences include the use of amber strobes instead of clear strobes in order to differentiate between fire notification and other life threatening events such as a bomb threat.

Traditional notification (horn strobes) or Voice Evacuation systems may be grouped together and accessed/initiated by radio frequency (RF). Cooper Notification's WAVES® product line offers command and control capability that allows system operators to access traditional Wheelock® horns and horn strobes, or individual and multiple Safepath® systems from a PC-based command and control security point. Systems may be addressed by zone and prerecorded messages or live voice announcements can be initiated. The WAVES® product line also includes large outdoor high-power fixed and/or portable (TACWAVES) horn arrays that can be configured as standalone access points in the system—thereby achieving indoor as well as outdoor coverage. The outdoor horn arrays can be powered by local ac power, or solar charged batteries—eliminating cable installations in large geographical land mass situations, Fig. 40-13.

Mass Notification may also include network alerting, such as blast e-mails to networked PCs and text alerting over cellphones and Blackberries®, as well as alerting pocket pagers.

Cooper Notification's Roam Secure® product line may be tied into the entire system for a complete notification solution.

40.5.1.1 Case Study

Scenario. An incident indicating universal alerting with as much coverage as possible, such as a random act of violence on a college campus.

Solution. Mass Notification. Campus security with one action can initiate live or prerecorded messages with flashing strobes over targeted zone loudspeaker systems inside and outside the buildings, while simultaneously broadcasting text messages over LAN networks as well as cell phones and Blackberries®.

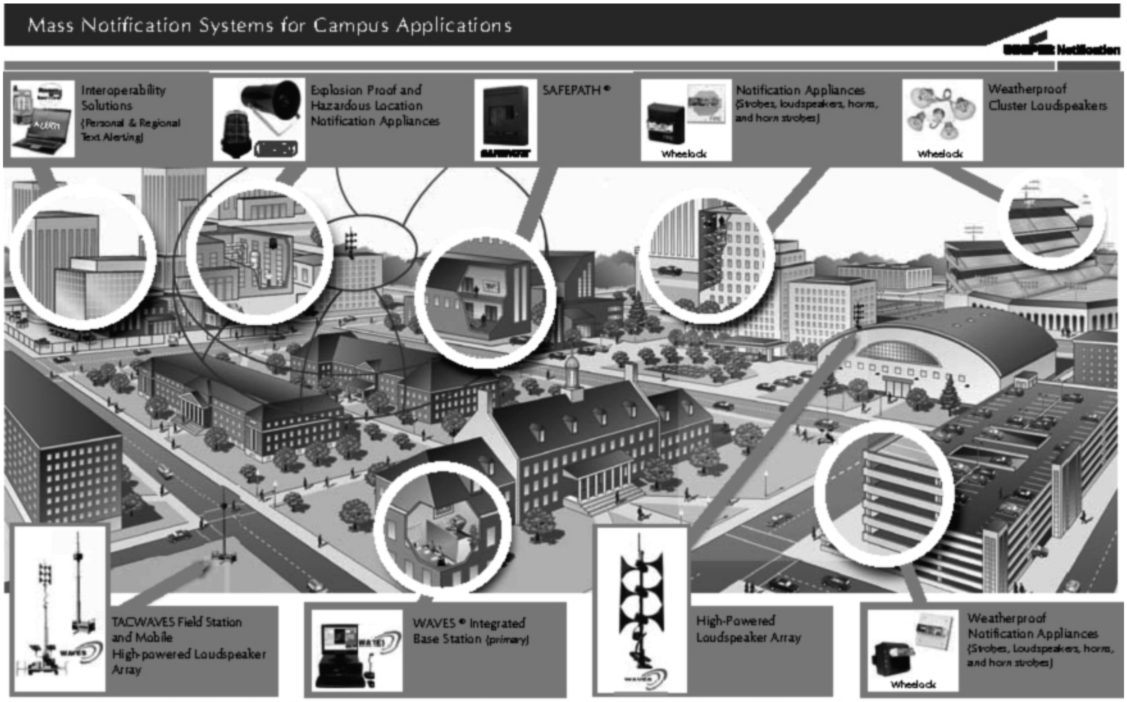

Mass Notification Systems for Campus Applications
COOPER Notification
Interoperability Solutions (Personal & Regional Text Alerting)
ALERT
Explosion Proof and Hazardous Location Notification Appliances
SAFEPATH ®
Notification Appliances (Strobes, loudspeakers, horns, and horn strobes)
Wheelock
Weatherproof Cluster Loudspeakers
Wheelock
TACWAVES Field Station and Mobile High-powered Loudspeaker Array
WAVES
WAVES ® Integrated Base Station (primary)
High-Powered Loudspeaker Array
Weatherproof Notification Appliances (Strobes, Loudspeakers, horns, and horn strobes)
Wheelock

Figure 40-13. Life Safety/Voice Evacuation system. Courtesy Cooper Notification.

40.6 Audio Archival

In addition to message repeating applications, there is a need for long term storage—i.e., to preserve audio and video data for future generations. It has been only the last quarter century that we came out of the dark ages and into the domain of digital storage, Table 40-1.

As the world is changing from analog to the digital domain, the media for archiving must be improved. After all, the medium used 20,000 years ago for written information can still be read today, while today's media for audio storage lasts only a few years and playback equipment becomes obsolete.

Today, most memory for audio storage is accomplished through solid state technologies. Solid state technologies fall into four broad categories:

1. Electrical memories based on semiconductor IC technology.
2. Magnetic memories based on magnetic materials.
3. Optical memories based on the interaction of light with matter.
4. Molecular, chemical, or biological memory based on changes in the atomic, molecular, or biological level.

Table 40-1. Evolution of Audio/Visual Information Storage. Courtesy DIGIPRESS

Medium	Origin (year)	Coding	Access	Lifetime (years)
Drawing/text:	20,000 BC	analog	–	100,000
Stone	3000 BC	analog	–	5000
Papyrus	200 BC	analog	–	2000
Parchment	105 AD	analog	–	1000
Ancient paper	1800 AD	analog	–	50
Modern paper	1900 AD	analog	–	50
Microfilm	1948 AD	digital	projector	3
Magnetic tape			computer	
Still image:	1835 AD	analog		50
B&W film	1869 AD	analog	player	20
Color film			player	
Sound:	1877 AD	analog		100
Cylinder	1887 AD	analog	player	50
Mechanical disc	1935 AD	analog	player	3
Magnetic tape	1985 AD	digital	player	10
Optical disc			player	
Moving images:	1895 AD	analog		50
B&W film	1935 AD	analog	projector	20
Color film	1951 AD	analog	projector	3
Magnetic tape	1946 AD	analog	player	
Computer data magnetic tape	1948 AD	digital	computer	3

Electrical memory is the most used technology for digital audio storage. Information is stored in digital form in various types of memory units. Common circuits are DRAMs, PROMs, EPROMs, flash EEPROMs, and ROMs.

Dynamic random access memory devices (DRAMs) store information dynamically, that is, as a charge on a capacitor. These designs feature one field-effect transistor (FET) to assess information for both reading and writing and a thin-film capacitor for information storage. Most nonvolatile cells rely on trapped charge stored on the floating gate of the FET. These units can be rewritten many times, the limit being determined by programming stress-induced degradation of the dielectric. Erasure of the charge from the floating grid is accomplished by tunneling or by exposure to ultraviolet light.

DRAMs are volatile, the average memory is about 10 years. Programmable memories can be programmed at least once and some can be programmed a million times. A few nonvolatile memories are programmable just once. These have an array of diodes or transistors with fuses or antifuses in series with each semiconductor cross point.

Electrically programmable read only memory devices (EPROMs) are usually used to describe cells that are electronically written and UV erased. EEPROM is probably the most common technology used. Static random access memory devices (SRAMs) are sometimes connected to EEPROM for storage when power is removed. Flash EPROMs require bulk erasure and therefore cannot be written over by the consumer.

Read only memory (ROM) is the only form of semiconductor storage that is permanently nonvolatile. Even with no power source present, information is retained in a ROM without any information loss.

Optical storage devices, the CD and DVD , are popular for long term archiving for the following reasons:

- Disk medium is highly standardized.
- Disk medium is multimedia (sound, data, still images, moving images).
- Disk medium format has a commercial life expectancy of many decades.
- Disk medium is an efficient and evolving medium.
- Disk medium has good chemical and mechanical resistance.
- Disk medium has good resistance to harsh environmental conditions.
- Disk medium has contactless reading—i.e., nondestructive.
- Disk medium is cost effective.
- Disk medium is an unrecordable system in the ROM version that prevents erasing or overwriting.

Archived CDs must be chemically stable, have good resistance against scratching, breaking, etc., and must be tolerant to extreme conditions of temperature, humidity, and electromagnetic fields. Some companies, such as DIGIPRESS, produce a stable CD. Rather than using a polycarbonate substrate, the CENTURY-DISC ARK from DIGIPRESS has a desalcanized etched tempered glass substrat, which is covered with titanium nitride—a very resistant metal. They can reach a lifetime of over 200 years—not forever maybe, but a great deal better than our present mediums can.

Acknowledgments

Many thanks to Jean Roche from MessageRepeaters.com for her inputs and editing of this chapter.

Additional Reading

Storage Technology Assessment: Final Report, Technical Report RE-0016. NML National Media Lab, St. Paul, MN, (612) 740-3670.

Chapter 41

Interpretation and Tour Group Systems

by Glen Ballou

41.1 Interpretation Systems 1531
  41.1.1. Central Control Unit 1531
  41.1.2. Interpreter's Booth 1531
  41.1.3. Hard Wired Systems 1533
  41.1.4. FM Interpretation Systems 1534
41.2 Multichannel FM Induction Loop 1535
41.3 Infrared Systems 1536
41.4 Tour Group Systems 1538
  41.4.1. FM Transmitter 1539
  41.4.2. FM Receiver 1540
  41.4.3. All-in-One System 1540

This page intentionally left blank

41.1 Interpretation Systems

As the world gets smaller and smaller, communications become increasingly important. Countries must talk to countries, businesses to businesses, and people to people. Only a few years ago, simultaneous interpretation systems were only found in places such as the United Nations and NATO. Today businesses are doing business with partners around the world, religious organizations have international meetings, schools are multilingual, and video and audio conferencing is common place.

Designing and building a simultaneous interpretation system is not just adding a set of earphones and another microphone to a sound system. A simultaneous interpretation system requires sound equipment and an acoustically correct room for the interpreter. The output from the various interpreters is transmitted to the various listeners in their language. This can be done via hardwire and earphones, AM or FM transmission, induction loop, or with infrared transmission systems.

Simultaneous interpretation systems allow a presentation by a talker to be heard and understood in or close to real time by all people in the audience. To accomplish this, the voice of the talker is directed to interpreters in soundproof booths or areas. The interpreters hear the original or floor language on headphones and instantly or simultaneously interpret it into the language they are assigned. The translated signal is then transmitted back into the audience area through the interpreters, microphones and transmission medium to the listeners through their control panel and headsets.

There are two basic types of simultaneous interpretation systems: bilingual and multilingual. Bilingual systems are designed for places where two and only two languages are used, such as in eastern Canada where French and English are used. Bilingual systems are the least expensive and the simplest to set up and use. These systems usually use only one interpreter's booth with either one or two interpreters.

Multilingual systems are used in the United Nations, large church conferences, boardrooms and schools, just to name a few. These systems are much more complicated and harder to install and use as they require individual interpreter rooms and a means for the listener to switch between languages.

41.1.1. Central Control Unit

The central control unit is the hub of the system. Most systems are microprocessor controlled and/or operated through an IBM compatible PC, Fig. 41-1. The floor language enters the unit at line level and is routed to the interpreters' booths and a tape recorder if required. The interpreted languages are returned to the central control from the interpreters' booths where they are prepared for transmitting to the listeners. This could be on hardwire, induction loop, infrared, or any combination of the three. Provision is also made for taping the interpreted language. The unit incorporates various operating modes and interlocks and a means for the interpreter and the operator to communicate with each other.

41.1.2. Interpreter's Booth

In the multilingual system, each booth normally has two or more interpreters that work as a team to interpret the floor language into the designated language of the booth. If many floor languages are allowed, each booth could require as many as four interpreters. The ISO standard for fixed interpreters' booths in systems with six to twelve languages recommends three interpreters per booth for the first six booths and four interpreters for the remaining booths, Fig. 41-2. Systems can have from two to thirty two languages, however, twelve seems to be the maximum normally used. Today most systems are digital, which can reduce background noise, distortion, and crosstalk. AGC assures equal listening level on all input channels, and the systems can be chained together with shielded FTP or STP Cat-5e cables, Fig. 41-3.

Booth size is specified by international standards. Permanent interpreters' booths and equipment are specified under ISO 2603 (1983), which specifies the minimum dimensions of 2.5 m wide × 2.3 m high × 2.4 m deep (8.2 ft × 7.75 ft × 7.87 ft). In booths with four interpreters, the width shall be 3.4 m (11 ft). An 80 cm (31.5 in) high window should extend the full width of the booth with the bottom of the window flush with the console. The room construction should attenuate the live sound so that if the nonreinforced sound does not exceed 80 dB, the inside signal will not exceed 35 dB.

Portable interpreters' booths are specified by ISO 4043 1981, and sound transmission using infrared is specified by IEC 764. The MB 2932 interpreter booth by Listen Technologies is intended for portable or fixed installations. The booth consists of four window panels, three blind panels, a door panel, a table, and two roof panels. It includes two ventilation fans and exceeds ISO 4043 sound insulation standards, Fig. 41-4.

Simultaneous interpretation systems are not just a simple input to an interpreter's booth. The system must enable the interpreter to hear the talker and to distribute

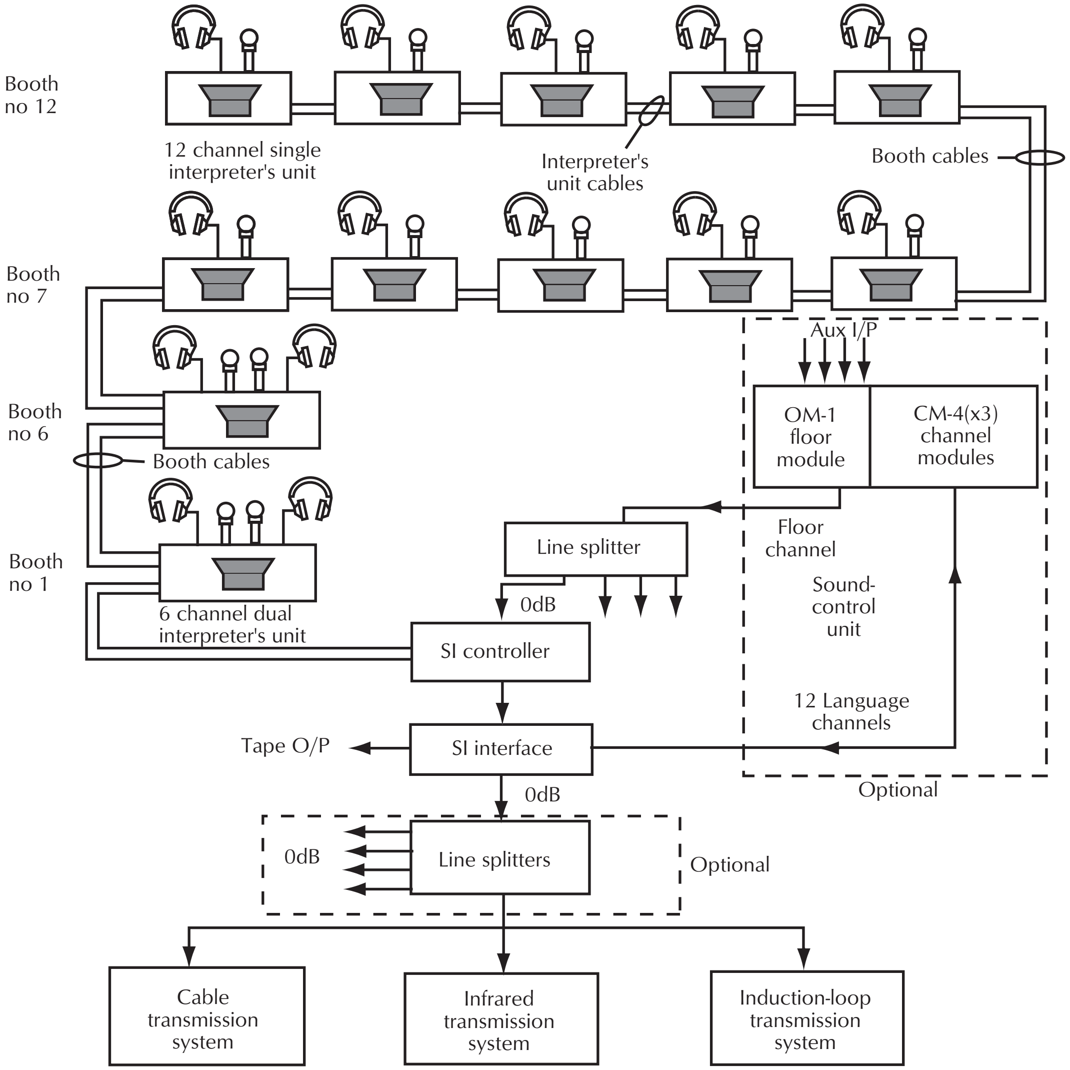

Booth no 12
12 channel single interpreter's unit
Interpreter's unit cables
Booth cables
Booth no 7
Booth no 6
Booth cables
Booth no 1
6 channel dual interpreter's unit
Aux I/P
OM-1 floor module
CM-4(x3) channel modules
Floor channel
Sound-control unit
12 Language channels
Optional
Line splitter
0dB
SI controller
SI interface
Tape O/P
0dB
0dB
Line splitters
Optional
Cable transmission system
Infrared transmission system
Induction-loop transmission system

Figure 41-1. Interpretation system. Courtesy Auditel Systems Limited.

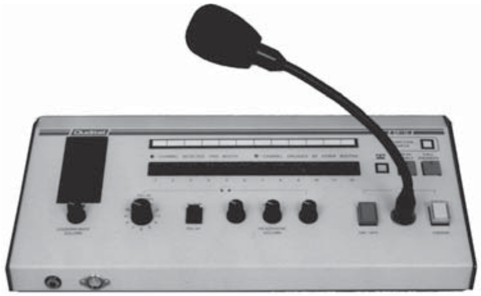


Figure 41-2. Twelve language interpreter's unit. Courtesy Auditel Systems Limited.

the interpreted language back to the language distribution system where it is sent to the various listeners. Beyond this the system must accommodate the following operations:

- One of the most important operations is to route the floor language through any unengaged channels. This is used so whenever the floor language is the same as the interpreted language of that particular booth, the interpreter does not have to interpret or remove his headset or change channels. This gives the interpreter a break.

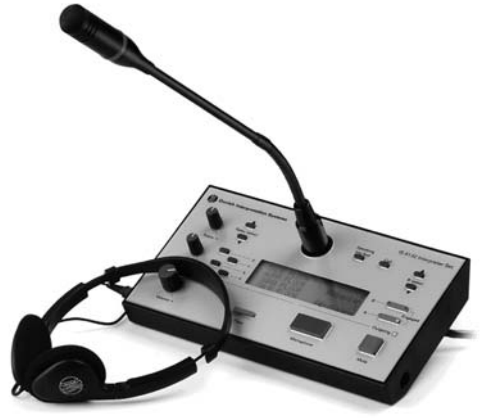


Figure 41-3. A fully digital two outgoing A/B channels and four input relay buttons. Courtesy Listen Technologies Corporation.

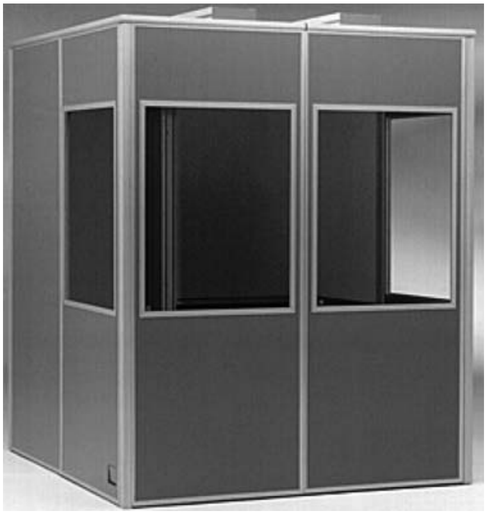


Figure 41-4. A portable or fixed installation interpreter's booth for two interpreters. Courtesy Listen Technologies Corporation.

- In the event the interpreter cannot understand the floor language, the console must have a *relay* facility so the interpreter can select one of the interpreted channels that she can understand, and make an indirect interpretation of that language into the booth designated language.
- It is advisable to have some degree of flexibility and control over the output channel of the booth to make fullest use of the capabilities of the team of interpreters in the booth. There must also be a means to prevent inadvertent interference with an engaged channel by another booth.
- When necessary, the output of any or all channels must be available for recording, and/or connecting to other feeds for transmission elsewhere.
- A final important accessory to the system is visual indicators on the console to identify engaged channels, and a two-way visual/audio system between the system operator and the interpreters' booth to summon the operator for assistance or alert the interpreter of a problem.

Figure 41-5 is an interpreter terminal for two interpreters according to ISO 4043. The unit is a double interpreter terminal for alternating operation and includes two microphone/headphone combinations. The two output channels are directly selectable by the interpreter and relay translations are possible from all languages. The listening area contains volume, bass, treble controls, an incoming channel selector, and an original/relay lever switch. It also has extra communication channels to and from the system operator and status information lights.

Once the language has been interpreted and sent to the master station, it must be routed to the listeners. There are three basic systems of transmitting the signal; the hardwired system, the multichannel FM inductive loop, and the infrared transmission system.

41.1.3. Hard Wired Systems

Hardwired systems are primarily used to transmit interpreted language channels to delegate stations on the conference hall floor. They are most useful in areas such as the United Nations building where the listeners are always seated in the same place and can tolerate the cable to the earphones. A hardwired system is the most reliable, has the best security against eavesdropping and has the best audio performance. As a rule, hardwired systems are cheaper in hardware costs but more expensive in installation costs. In multiconductor cable systems, each channel is amplified and transmitted on a pair of conductors. Each listener usually has a panel located at his or her seat that includes a language selecting switch, a volume control, and an earphone jack. If the conductors are a twisted pair, there is little crosstalk in lines in excess of 1000 m (3280 ft), and farther with shielded cable. Hardwired systems are not particularly good for portable systems as it is not easy or physically safe to lay out cables on the floor to the various listeners. It is important that the user does not place the earphones next to a microphone, unless it is

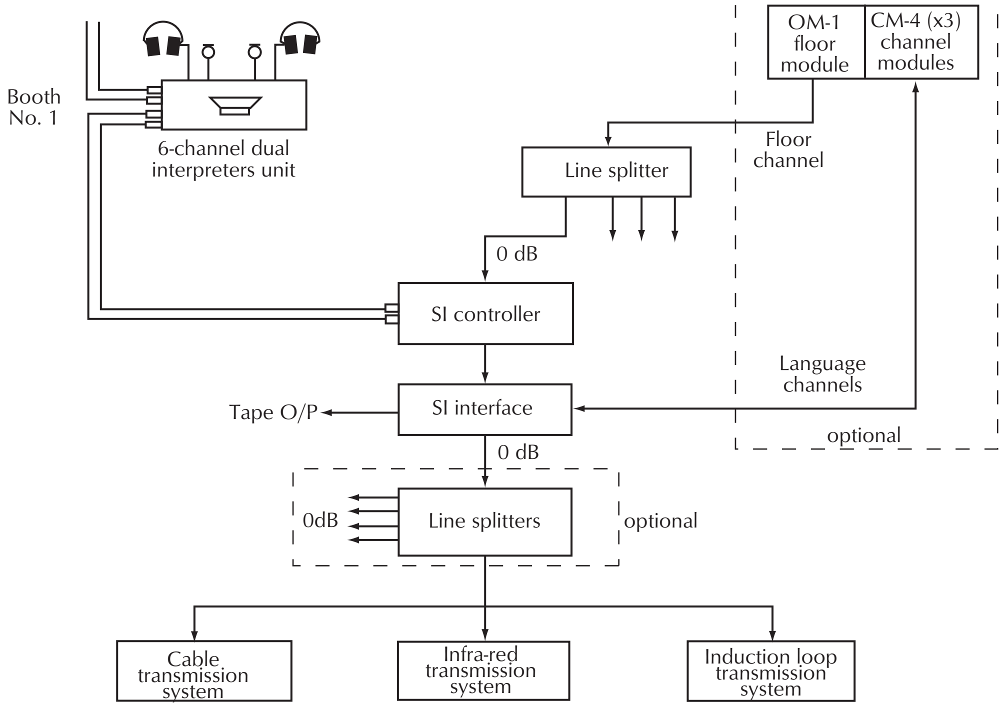

OM-1
floor
module
CM-4 (x3)
channel
modules
Booth
No. 1
6-channel dual
interpreters unit
Floor
channel
Line splitter
0 dB
SI controller
Language
channels
Tape O/P
SI interface
optional
0 dB
0dB
Line splitters
optional
Cable
transmission
system
Infra-red
transmission
system
Induction loop
transmission
system

Figure 41-5. Two person interpreter's terminal.

shutoff, as they may be on the same channel and cause feedback. Acoustical crosstalk can occur if open back earphones are used because an adjacent live microphone can sometimes pick up the interpreted language.

If many languages are used, it may be better to multiplex the signal rather than use a multiconductor cable. This system would consist of a central modulator with up to twelve channels driving a network of active channel selector units using coaxial cable in a loop-through configuration. Power for the channel selectors is provided by power supplies injecting dc into the network.

Most single cable conference systems with delegate microphone units incorporate a built-in loudspeaker. The loudspeaker signal is derived directly from a common audio line. This simplifies cabling by avoiding the necessity for a second audio line to drive the loudspeakers. This does mean that the input and output signal are on the same line, and would create a closed loop and feedback unless some means of isolating the two signals is employed. This problem is overcome by Auditel with the application of a common mode reverse audio feed (CMRAF), Fig. 41-6. The technique is based on selective rejection of large common mode signals. The output from the microphone preamplifier is a balanced signal and is extracted in the central unit via transformer. After signal processing, the loudspeaker drive signal is injected into the audio pair in common mode form. Since the loudspeaker drive amplifiers and delegate units reject balanced signals the two signals can be carried over the same conductors without interaction or without compromising the signal quality.

41.1.4. FM Interpretation Systems

FM products can be used for language interpretation by connecting a stationary FM transmitter to an audio system transmitting an FM signal to a portable receiver for assistive listening and a language interpreter. In addition to the portable receiver, the interpreters use a portable transmitter and an over-the-head microphone and earphone unit. This combination allows them to hear the audio clearly in an adjoining area while speaking their translations in a normal tone of voice. Their translations are sent via FM back to participants' receivers. It is important to have transmitters and receivers with multiple channels allowing users to find clear channels even in a crowded venue with extensive FM use.

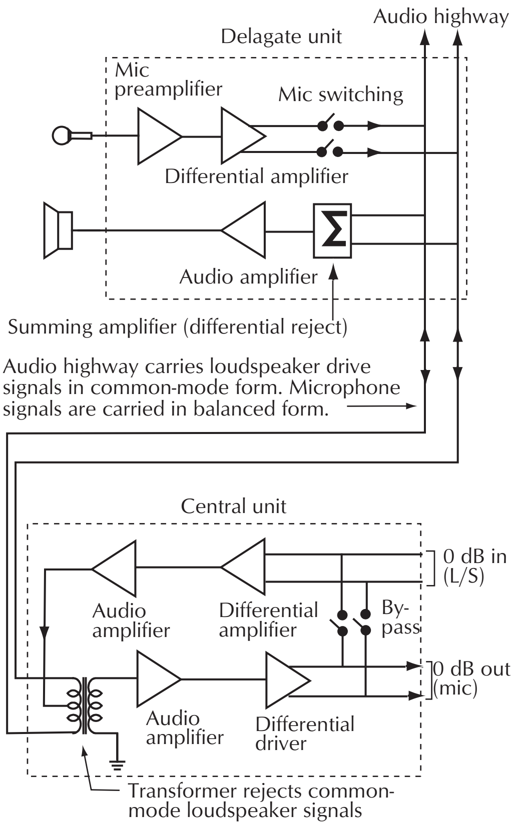

Audio highway
Delagate unit
Mic preamplifier
Mic switching
Differential amplifier
Σ
Audio amplifier
Summing amplifier (differential reject)
Audio highway carries loudspeaker drive signals in common-mode form. Microphone signals are carried in balanced form.
Central unit
0 dB in (L/S)
Audio amplifier
Differential amplifier
By-pass
0 dB out (mic)
Audio amplifier
Differential driver
Transformer rejects common-mode loudspeaker signals

Figure 41-6. Common mode audio. Courtesy Auditel Systems Limited.

41.2 Multichannel FM Induction Loop

The induction loop system is a development of the audio loop for assisted hearing that has been around for many years. This system has the advantage of being less expensive than the infrared system and more portable than the hardwired system. It is not affected by line-of-sight limitations and does not require visible radiator panels as is required by infrared systems. Originally these systems operated on the AM band, however, today's induction loop systems operate on the FM band for better quality and less noise and interference.

A closed loop antenna is installed around the perimeter of the area if the room is small, 30 ft × 60 ft (9 m × 18 m), Fig. 41-7A, or as a zigzag or circular pattern in larger areas with a 10 ft × 15 ft (3 m × 4.5 m) pitch Figs. 41-7B and 7C. The zigzag or circular system should always be on or in the floor rather than in the ceiling because the field strength above and below the loop is reduced as the horizontal strength is increased. The antenna should be placed in aisles where pickup is not essential.

The loop consists of a single conductor with a minimum cross section area of 2.5 mm^2. Loop inductance should be about 1.5 μH/m. This translates to 1.5 mH for a 1000 m loop. For best signal, the FM transmitter should be capable of delivering 100 mArms per channel to the antenna. It is common to install the loop on or in the floor rather than in the ceiling because the seated listener is usually closer to the floor, and ceilings are apt to have a lot of metal that shorts out the signal. If the antenna is installed permanently in conduit in the floor, nonmetallic conduit, such as PVC, must be used.

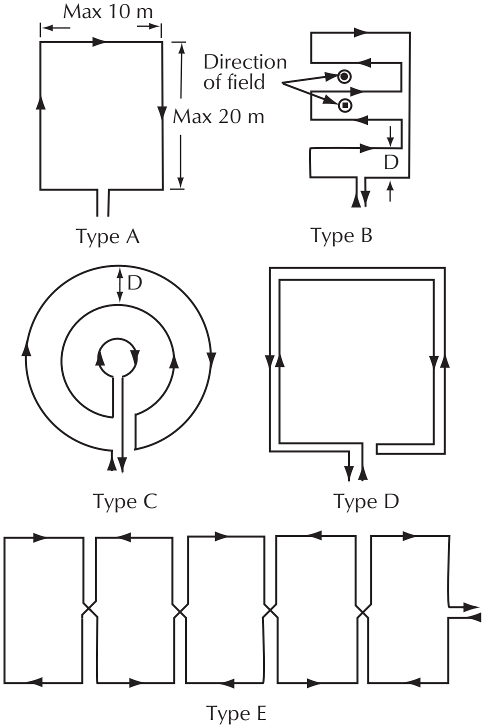

Max 10 m
Max 20 m
Direction of field
D
Type A
Type B
D
Type C
Type D
Type E

Figure 41-7. Configuration for inductive loops. Courtesy Auditel Systems Limited.

An FM modulated-carrier transmits the signal. The band for the carrier frequencies is limited to 15–150 kHz by international Telecom regulations in most areas except North America. This bandwidth limits the number of voice channels to about eight. The radiation is primarily magnetic and pretty much confined to within the area defined by the loop, therefore the system has reasonable security. Audio quality is

Table 41-1. Technical Specifications for FM Inductive Loop Systems

Channel No	1	2	3	4	5	6	7	8
Channel Frequencies (kHz)	59.7	44.8	29.9	134.3	119.4	104.5	89.6	74.6
Modulation	FM							
Normal Deviation	±1.5 kHz							
Peak Deviation	±1.7 kHz							

Audio Frequency Characteristics	Transmitter	Receiver
Frequency response (–3 dB)	125 Hz–4 kHz	90 Hz–4 kHz
Max. distortion (at 1 kHz)	<1%	<2.5%
Signal/Noise (A weighting)	>50 dB	>45 dB

restricted to a 3–4 kHz bandwidth that is adequate for voice only. Table 41-1 gives the technical specifications for an eight channel inductive loop system by Auditel.

Inductive loop systems are inexpensive and easy to install and therefore usually preferred in large portable and multiple venues where cost is important. Different systems in adjacent areas can cause crosstalk and security can become a problem if a listener can get close to the loop. Steps can be taken, such as using an out-of-phase loop around the room about 3 ft beyond the periphery of the transmission loop, to reduce horizontal coverage, Fig. 41-7D. To reduce vertical interference where one system is directly above another, a double zigzag antenna system can be used, Fig. 41-7E. This antenna system will reduce the unwanted signal to 50 dB below the wanted signal when the separation distance is 10 ft (3 m) as opposed to 20 ft (6 m) with a standard antenna.

The induction loop receivers normally operate on batteries so the listener can be located or move anywhere within the induction loop. The receivers use a ferrite rod antenna, have a battery status indicator, a volume control, and a channel selector switch. Most units use alkaline batteries and get upwards of 500 h between battery changes. High fidelity headsets are not required because the overall frequency response is so limited.

41.3 Infrared Systems

An infrared system is a modern version of the wireless system. InfraRed was designed by Sennheiser in the seventies as a medium for wireless audio transmission. It was initially used as a broad band system for home entertainment and later for assisted hearing systems. It is applied to multichannel audio systems using narrow band modulation techniques in the 930 nanometer wavelength band. Rather than using radio frequencies for transmitting the signal, it uses infrared frequencies, which are confined to line-of-sight or reflections off of objects. These objects can be mirrors, glass ashtrays, or any other brightwork. Infrared goes through glass windows and is absorbed by objects and walls that are dark green or black. It is important that the walls of a room using infrared be light colored and have a minimum number of windows.

Two different transmission techniques are used: the pulse modulated system and the FM multiplexed system. The pulse modulated system has an advantage in that the amount of emitted radiation required is independent of the number of channels, however, it usually has poorer audio quality than the FM multiplexed system. Pulsed systems do not meet IEC audio standards and can be affected more by high-frequency fluorescent lights.

Most systems used today are based on modulated carrier techniques using FM. The operating frequencies for wide-band two-channel infrared systems are 95 kHz and 250 kHz with peak deviation of ±50 kHz. Narrow-band systems operate on twelve or more channels between 55 kHz and 1335 kHz (excluding 455 kHz) with 40 kHz channel spacing and peak deviation of ±7 kHz. These standards are specified in the IEC 76 international standard and insure compatibility between manufacturers. General specifications for infrared systems are given in Table 41-2. The system comprises three sections: the transmitter, the emitter (sometimes both are combined in one unit), and the receiver. The transmitter imparts the audio signal onto a subcarrier that the emitter converts into infrared light. The receiver decodes the infrared signal to retrieve the original audio, Fig. 41-8.

The signal enters the listening area via infrared radiators. The IR light emitting diode can cover an area of 70 ft^2 and has a coverage angle of ±25 degrees. The diode incorporates a collecting lens which bends the infrared light to hit the diode, Fig. 41-9. Increasing the number of diodes increases the IR intensity and the area of coverage multiplies by the number of diodes in the panel. The diodes have a continuous operating life of 100,000 h before the light output diminishes to 70% of its original value.

Table 41-2. Technical Specifications for Infrared Systems

Characteristics	Narrow Band System	Wide Band System
No. of channels:	12	2
Carrier frequencies:	55 kHz–535 kHz (excl. 455 kHz)	95 kHz and 250 kHz
Channel spacing:	40 kHz	155 kHz
Modulation:	FM	FM
Pre-emphasis:	100 µs	50 µs
Normal deviation:	±6 kHz	±35 kHz
Peak deviation:	±7 kHz	±50 kHz
Transmitters:		
Frequency response (–3 dB):	50 Hz–8 kHz	50 Hz–13 kHz
Max. distortion at 1 kHz:	<1.0%	<1.0%
Signal/noise (A weighting):	>55 dB	>70 dB
Receivers:		
Frequency response (–3 dB):	50 Hz–8 kHz	100 Hz–9 kHz
Max. distortion (at 1 kHz):	<2.5%	<1.0%
SNR (A weighting):	>55 dB	>63 dB
Emitter Panels:		
Frequency response (–3dB):	30 Hz–710 kHz	

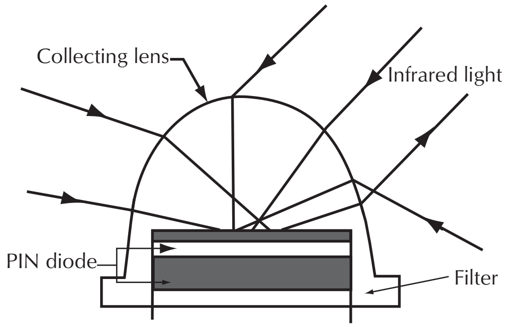
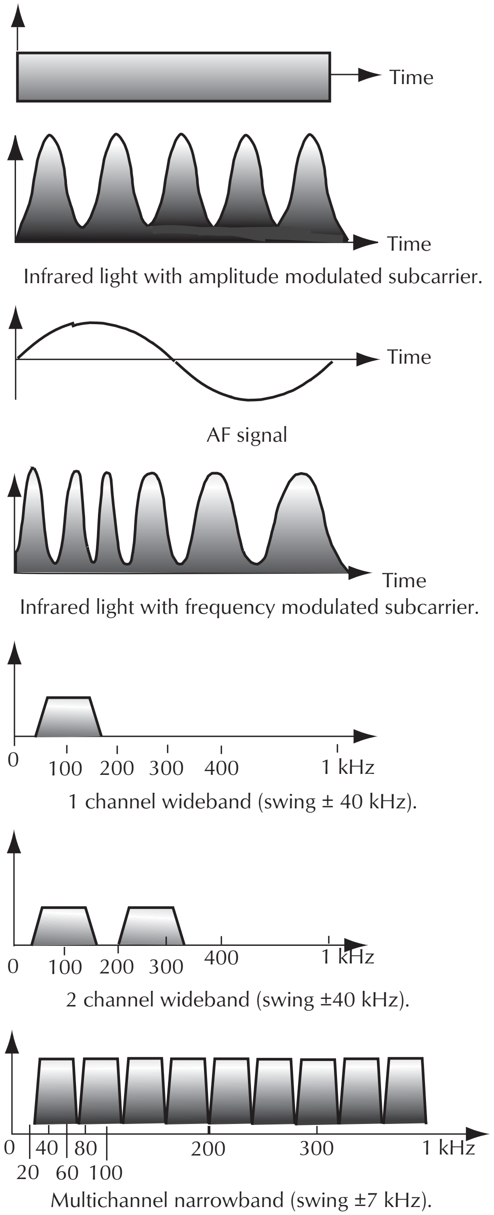


Figure 41-8. Light emitting diode (LED). Courtesy Sennheiser Electronics.



Infrared light with amplitude modulated subcarrier.

AF signal

Infrared light with frequency modulated subcarrier.

1 channel wideband (swing ± 40 kHz).

2 channel wideband (swing ±40 kHz).

Multichannel narrowband (swing ±7 kHz).

Figure 41-9. Infrared transmission characteristics.

Depending on the size of the room, its shape, and surface characteristics, a single small radiator or multiple large radiators may be required. Auditel Systems Limited states that the number of radiators required for a SNR of >40 dB can be calculated using the following equation:

$$N = \frac{area(m^2) \times number\ of\ channels}{D} \qquad (41\text{-}1)$$

where,

$$D = \frac{\textit{total emitted power}\ (\mathrm{mW})}{\textit{receiver sensitivity}\ (\mathrm{mW/m^2/channel})}.$$

This does not take into account the wall surfaces, niches, and obstructions and it assumes that at least 95% of the radiation is usable. Sennheiser states its large radiator can cover 11,000 ft^2/number of channels in the system.

Layout of the panels is also important. Every seat must have a view of a panel. The range and coverage of a radiator are influenced by its orientation to the surface to be illuminated. A panel that is located so its pattern is parallel to the floor will have a long footprint with decreasing signal with distance. A panel that is aimed straight down will have a circular pattern that has about the same signal everywhere.

Total infrared power is proportional to the number of diodes in the panel. Power can be doubled by using two panels at the same location or using a panel with twice as many diodes. It is often best to use more than one radiator to eliminate dead spots in the area. If the overlapping areas have reduced signal, the two signals can add to bring the SNR ratio up to an acceptable level.

The infrared signal is received by the listener through a belt pack or integral headset type of receiver. The body pack receiver is about 125 mm × 60 mm × 28 mm (4.9 inches × 2.3 inches × 1.1 inches) and weighs 100 g (3.5 oz) and incorporates a channel selector switch, volume control, and earphone jack. The integral headset receiver hangs from the ears down below the wearer's chin. It includes a channel selector and volume control, weighs 2.1 oz, and can deliver 110 dB, Fig. 41-10.

Security from eavesdropping is good since an unwanted receiver cannot see the transmitter or its reflections. Interference between rooms is also good since infrared cannot go through solid walls. Infrared systems are affected by other infrared sources such as sunlight and incandescent and fluorescent lights. Infrared is never usable in direct sunlight but can often be used in shaded areas if high-power transmitters are used. Rooms with intense incandescent and fluorescent lighting may require high-power transmitters. Because it is line-of-sight, objects, including people between the transmitter and the receiver, can cause dropout unless at least two transmitters are covering the same area.

Infrared systems are used for portable systems, where a large room can be subdivided, and for fixed installations. This system is more expensive than the induction loop but has much better audio quality, and is not as susceptible to electrical interference. Infrared is the system of choice today.

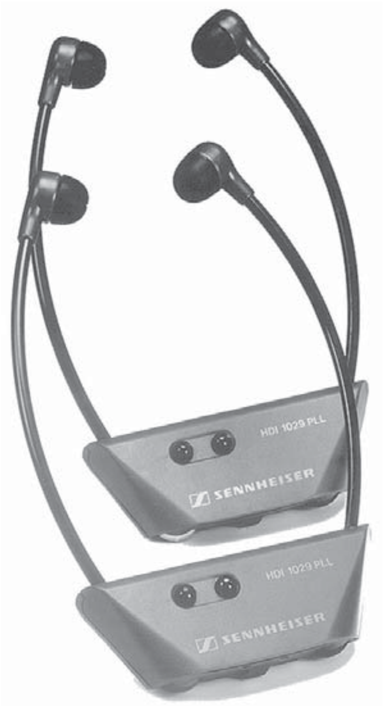


Figure 41-10. Infrared headset with built-in receiver. Courtesy Sennheiser Electronics.

41.4 Tour Group Systems

A tour group FM system consists of a portable FM transmitter and portable FM receivers. A microphone, often worn over the head, is connected to the transmitter and broadcasts the presenter's voice to everyone in the audience or group. The portable transmitter allows the audio to be delivered without having to carry a microphone or be plugged into the wall. Participants wear a portable FM receiver with an earphone to hear the presentation. An unlimited number of receivers can be used with one transmitter as long as the participants are within the broadcast range, typically up to 150 ft (45.7 m).

The transmitter and the receivers are tuned to the same channel and depending on the frequency, three–eight channels can be used simultaneously. This allows for multiple tours to be conducted at a time and/or provide language interpretation. Currently tour group products are available in 72 MHz, 216 MHz, and 863 MHz.

Transmitters and receivers offer a mix of features and functionality for channel selection, programming, power, signal strength, and use. Portable transmitters and receivers are typically battery powered with standard alkaline or NiMH batteries.

Tour group systems are excellent for factory tours, museums, outdoor events, wireless microphone applications, classroom or training, or personal use. Anywhere you need to amplify sound but don't have (or want) an installed sound system.

41.4.1. FM Transmitter

A typical portable FM transmitter is shown in Fig. 41-11. The specifications for this transmitter are:

General

- Number of channels: 19 wide-band, 38 narrow-band.
- Channel tuning capable of being locked.
- SNR: 70 dB or greater.
- Output power: adjustable to quarter, half, or full.
- Audio frequency response: 50 Hz–15 kHz ±3 dB.
- Includes: a microphone sensitivity switch.
- Includes: a mute switch.
- Operates on two AA batteries.
- Includes: an LCD display that indicates battery level, channel, channel lock, low battery, battery charging, programming, and RF signal strength.
- Includes: automatic battery charging circuitry for recharging NiMH batteries.

Radio Frequency

- RF frequency range: 216.0125–216.9875 MHz.
- Frequency accuracy: ±0.005% stability from 32 to 122°F.
- Transmitter stability: 50 PPM
- Transmitter range: 0 to 150 ft (45.7 m).
- Output power: less than 100 mW (216 MHz).
- Antenna: the microphone cable.
- Compliance: FCC Part 15, Industry Canada.

Audio

- System frequency response: 50 Hz–10 kHz ±3 dB 216 MHz.
- SNR: SQ enabled; 80 dB; SQ disabled 60 dB.
- System distortion: <2% harmonic distortion (THD) at 80% deviation.
- Microphone input: unbalanced, +4 dBu maximum, –10 dB nominal input level adjustable, impedance 10 kΩ).
- Microphone sensitivity: three position switch: high, middle, and low in 6 dB increments.
- Line input: unbalanced, –10 dBu nominal input level, –3 dBu maximum, impedance 10 kΩ.
- Microphone power: 3 Vdc bias.

Controls

- User controls: power, mute, channel up/down.
- Setup controls: in the battery compartment, microphone sensitivity, NiMH/alkaline battery, SQ enable/disable.
- Programming: channel lockout, channel lock.

Indicators

- LED red: illuminated when unit is on, flashes when batteries are low, or to indicate charging. Flashes two times when muted.
- Display: channel designation, lock status, signal strength indication, battery life, RF power.

Power

- Battery type: 2 AA batteries, Alkaline or NiMH.
- Battery life: alkaline –10 h, NiMH rechargeable.
- Battery charging: (NiMH only), fully automatic, 13 h.
- Power supply compliance: RoHS, WEEE, UL, PSE, CE, CUL, TUV, CB compliant.

Physical

- Dimensions: (H × W × D) 5.0 inches × 3.0 inches × 1.0 in (13.0 cm × 7.6 cm × 2.5 cm).
- Color: dark gray with white silk screening.
- Unit weight: 3.9 oz (111 g).
- Unit weight with batteries: 5.8 oz (164 g).

Environmental

- Temperature—operation: 14 to 104°F.
- Temperature—storage: –4 to 122°F.
- Humidity: 0 to 95% relative humidity, noncondensing.

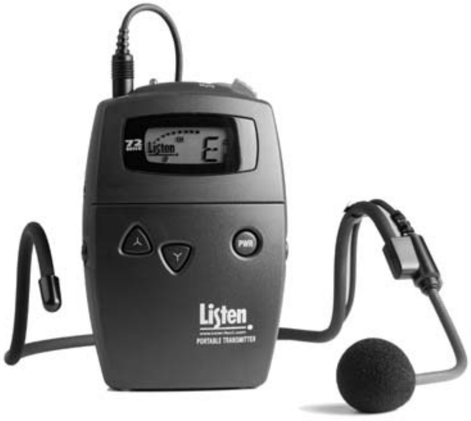


Figure 41-11. Typical high-quality FM transmitter, the Listen LT-700-863. Courtesy Listen Technologies Corporation.

41.4.2. FM Receiver

A typical FM receiver is shown in Fig. 41-12. The specifications for this FM receiver are:

General

- Number of channels: 19 wide band, 38 narrow band.
- SNR: 70 dB or greater.
- Programmable to electronically lock out unneeded channels.
- Can seek channels and can be locked on a single channel.
- Adjustable squelch.
- Audio frequency response: 50 Hz–15 KHz ±3 dB.
- Includes: a stereo headset jack for either a mono or stereo headset.
- Includes: an LCD display that indicates channel, battery level, low battery, battery charging, and RF signal strength.
- Functions in both DX and Local mode.
- Operates on two AA batteries.
- Includes: an automatic battery charging circuitry for recharging of NiMH batteries.

Radio Frequency

- RF frequency range: 216.0125–216.9875 MHz.
- Number of channels: 19 wide band, 38 narrow band.
- Sensitivity: 0.6 uV typical, 1 uV maximum for 12 dB sinad.
- Frequency accuracy: ±0.005 stability, 32 to 122°F.
- Antenna: uses earphone cable.
- Squelch: programmable in twenty steps, automatic on loss of RF signal.
- Compliance: FCC Part 90, Industry Canada.

Audio

- Frequency response: 50 Hz–10 kHz ±3 dB 216 MHz.
- SNR (A-weighted): SQ enabled: 80 dB; SQ disabled 60 dB.
- System distortion: <2% total harmonic distortion (*THD*) at 80% deviation.
- Output: unbalanced, 0 dBu nominal output level, 16 mW maximum, impedance 32 Ω.

Controls

- User controls: channel up/down, seek, volume.
- Set up controls (battery compartment): alkaline/NiMH batteries, SQ enable/disable.
- Programming: channel lock, squelch, channel lockout.

Indicators

- Red LED: illuminated when unit is on. Flashes when batteries are low, or to indicate charging. Flashes when locked and seek is pushed.
- Display: channel designation, lock status, signal strength indication, programming.

Power

- Battery type: two AA batteries, alkaline or NiMH.
- Battery life: alkaline–15 h NiMH rechargeable.
- Battery charging (NiMH only): fully automatic, 14 h.
- Power supply: Input—120 Vac, Output—7.5 Vdc 250 mA.

Physical

- Dimensions (H × W × D) 5.0 × 3.0 × 1.0 inches (13.0 × 7.6 × 2.5 cm).
- Color: dark gray with white silk screening.
- Unit weight: 3.9 oz (111 g).
- Unit weight with batteries: 5.8 oz (164 g).

The above specifications are fairly common for high-quality receivers.

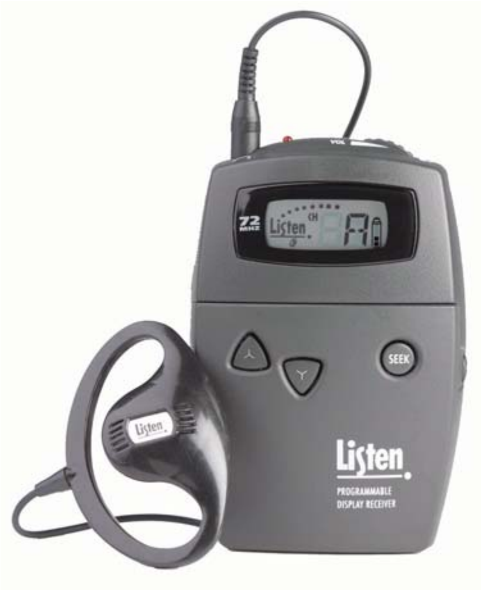


Figure 41-12. Typical high-quality FM receiver, the Listen LR-500-863. Courtesy Listen Technologies Corporation.

41.4.3. All-in-One System

An all-in-one system or self-contained system for use in multiple venues is Listen Technologies' Soundfield FM,

Fig. 41-13. Soundfield FM applications can be in a classroom, meeting room, training center, or theater. Studies have shown that boosting the audio levels of normal voices increases retention by as much as 30% for the listeners, and can reduce potential for voice strain and fatigue for the presenter.

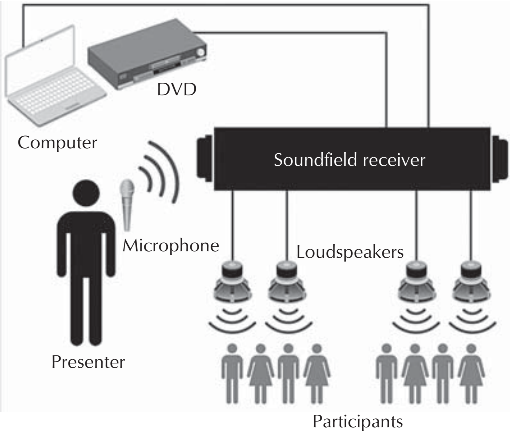


Figure 41-13. A portable system utilizing FM transmission for the microphone input. Courtesy of Listen Technologies Corporation.

Wherever a large PA system is impractical, a Soundfield FM-type solution is a possible answer. A Listen Soundfield FM system is an easy way to deliver interference-free sound to groups of all sizes insuring that all participants hear the message clearly.

The LR-100 Stationary Receiver/Power Amplifier, when used with a Listen transmitter, delivers audio for use in a variety of locations and applications—most commonly soundfields. The unit has an LCD display showing channel, lock, and battery level. A security cover protects auxiliary and receiver volume and trim controls. It has an adjustable squelch control, balanced or unbalanced loudspeaker wire output options, and multiple input options for flexibility. The system is powered by a 16 Vac/1000 mA power input or a12 Vdc battery. An antenna can be added for longer range applications.

The LR-600 is a wireless receiver, two-channel 10 W power amplifier and loudspeaker. The unit is powered by alkaline or NiMH rechargeable AA batteries, or a 15 Vac power supply or 12 Vdc, for 15–10 W continuous power output. It has unbalanced auxiliary input and output ports. The LCD display shows channel, lock, and battery level.

This page intentionally left blank

Chapter 42

Assistive Listening Systems

by Glen Ballou

42.1 Nature of the Problem 1545
42.2 Types of Assistive Listening Systems 1545
  42.2.1 Magnetic Induction Loops. 1546
    42.2.1.1 Loop Design Criteria 1546
  42.2.2 Loop Location and Size. 1547
    42.2.2.1 Loop Current 1548
    42.2.2.2 Loop Impedance 1548
    42.2.2.3 Electronic System 1549
    42.2.2.4 Installation 1550
  42.2.3 FM Broadcast 1550
  42.2.4 AM Broadcast 1551
  42.2.5 Infrared 1552
    42.2.5.1 Components of an Infrared System 1553
    42.2.5.2 Categories 1553
    42.2.5.3 Coverage 1553
    42.2.5.4 Ambient Light 1553
    42.2.5.5 The infrared Link 1553
42.3 Receivers 1555
References 1557

This page intentionally left blank

42.1 Nature of the Problem

There are millions of people in the world (20 million in America alone) with hearing impairments for whom the acoustical and electronic systems described elsewhere in this book are inadequate. It is surprising that the special hearing needs of so large a group have been largely ignored for so long, especially when each of us faces the very real probability of joining that group through disease, trauma, or just by growing old.

According to the National Association of the Deaf, NAD, assistive listening systems (ALSs), sometimes called assistive listening devices (ALDs), are amplifiers that bring sound directly into the ear. They improve the speech-to-noise ratio by separating the sounds, particularly speech, that a person wants to hear from background noise.

Research indicates that people who are hard of hearing require a signal-to-noise ratio increase of about 15–25 dB in order to achieve the same level of understanding as people with normal hearing. An ALS allows them to achieve this gain for themselves without making it too loud for everyone else.

ALSs are used by people with various degrees of hearing loss, from mild to profound, including hearing aid users and those with cochlear implants, as well as those who use neither. ALSs are sometimes described as "binoculars for the ears" because they stretch hearing aids and cochlear implants, thus extending their reach and increasing their effectiveness.

ALSs address listening challenges by minimizing background noise, reducing the effect of distance between the sound source and person with hearing loss, and overriding poor acoustics such as echo. ALSs are used in places of entertainment, employment, education, and home/personal use.

The hearing impaired are not just the people who wear aids. In fact, only about 20% of the hearing impaired wear aids. Many people with hearing losses are able to function in close or face-to-face situations but are lost in noisy or reverberant settings. Even people who wear hearing aids have problems in reverberant rooms or where there is a high background noise level. Our standards for speech intelligibility are based on listening tests with normal-hearing subjects and are not directly applicable to the hearing impaired, see Chapters 2 and 40. Noise and reverberation degrade intelligibility far more rapidly for the hearing-impaired individuals whether they are fitted with hearing aids or not. Often the very highly prized acoustical qualities of theaters and concert halls operate against the needs of the hearing impaired, and the acoustical design of most classrooms and lecture halls are inadequate for the hearing-impaired student.

For years the only special assistance offered to the hearing impaired was headphones in a couple of pews in the front of the church sanctuary. In recent years new wireless technologies have been developed or adapted to meet the special needs of the hearing impaired in public assembly spaces. No longer is the user restricted to a specially wired seat; now every seat is available; no special ticketing is required. The user is free to sit with family and friends.

Wide area ALSs are covered under Title III of the ADA (Americans with Disabilities Act of 1990). This title stipulates that ALSs be provided in public places unless a provider can prove that it is an undue burden. Examples of such venues include movie cinemas, live performance theaters, and public classes. The ADA specifies that ALS receivers be provided at no cost and specifies the number of receivers that must be provided depending on the number of seats (4% rule). Revised ADA Guidelines to be released in the future are expected to increase standards for performance of ALS and address related issues.

ALSs may also be indicated under ADA Title I (employment accommodations) as well as Title II (accommodations provided by state and local governments). Other public policies that may require use of ALSs include Section 504 of the Rehabilitation Act (affecting federally funded agencies) and Individuals with Disabilities Education Act.

42.2 Types of Assistive Listening Systems

There are four basic types of wireless systems: magnetic induction, FM broadcast, AM broadcast, and infrared light. Each type has its own set of advantages, problems, and limitations. There is no single best system for every application; each system is simple to operate and to install.

The system, no matter what type, must pick up the program sound. In a fully mic'ed event, this pickup could be a feed from the reinforcement control console. Where the event is not mic'ed, there must be a special microphone or microphones to feed the hearing-impaired system. It is very important that the feed to the system be of the highest quality possible with a minimum of reverberation pickup and extraneous noises. A pressure zone-type microphone, see Chapter 16, on the forestage floor or mounted on an acoustical reflector panel over the forestage would be good for many shows. An even better system, which would reduce room effects, would be to individually or close-mic the actors, talkers, or singers. A sound pickup that does not reject the reverberant field and extraneous noise, and/or has distorted sound will not be useful for a system for the hearing impaired.

42.2.1 Magnetic Induction Loops

Magnetic induction, sometimes called a *loop system*, is one of the oldest but still useful systems. The principal advantage of this system is that it can operate directly into the user's hearing aid without the need for a portable receiver as required by all other systems. A loop of wire is wrapped around the seating area, usually under the carpet, and connected to an amplifier. The electrical current flowing through the loop will create a magnetic field (as the primary of a transformer) that can be picked up by a hearing aid equipped with a *T-coil* (T as in telephone). About 60% of the hearing aids in the United States have T-coils for magnetic coupling with the earpiece of a telephone. Portable receivers are available for use by patrons who do not have an aid with a T-coil.

There are, however, several problems with the loop system. Most buildings have other magnetic fields that will be picked up by the T-coil. Ordinary electrical wiring will radiate a large 60 Hz field throughout the room, so the T-coil and the portable receivers are designed to have no response in the low-bass region in order to avoid the 60 Hz hum. There are other power-line-related noises that cannot be filtered out—motors, dimmers, and fluorescent lamps being the most common. The size and shape of the loop and the amount of nearby steel in the building or in the seats will affect the strength and uniformity of the magnetic field. Simultaneous use of loops in adjoining rooms is often a problem because of crosstalk between the systems. Quality of reception is also dependent upon the quality of construction of the T-coil in the hearing aid. All these factors combine to provide reception that sometimes has poor sound quality, is noisy, and varies in volume depending on the location of the listeners and how they turn their heads.

The limited statistics available indicate that of the people needing hearing assistance, only about 20% are actually wearing aids, and only 60% of those aids are equipped with T-coils, which suggests that only 12% of those needing help are able to make use of a magnetic loop through their hearing aids. It has been argued that the majority of those without T-coils actually are young children and the very old; active adults are most likely to have T-coils. Despite the comparatively limited availability of T-coils and the several substantial limitations of the magnetic loop system, it remains popular and enjoys the vocal support of many of those 2.5 million people who have T-coils.

The cost of a magnetic induction loop system is largely the cost of the amplifier and the labor of installing the wire loop. The receivers are inexpensive. Until recently, the magnetic loop was the least costly system. However, advances in solid state electronics have made the AM and FM broadcast systems very competitive in price. Where large areas are to be covered, the magnetic loop is probably not as cost effective as a broadcast system.

42.2.1.1 Loop Design Criteria

The international standard for the magnetic field strength of a loop system with an input signal of normal speech level is 0.1 A/m. Magnetic field strength H = 0.1 A/m in SI units or 0.125 Oe in cgs units. This field strength produces an audio voltage in the T-coil about equal to the output of the hearing aid's microphone at normal speech levels, Fig. 42-1. This eliminates the user having to make volume control adjustments when switching between microphone and T-coil. Also, this field is strong enough that noise and interference problems are minimized, yet it is not so strong as to overload the hearing aid amplifier.

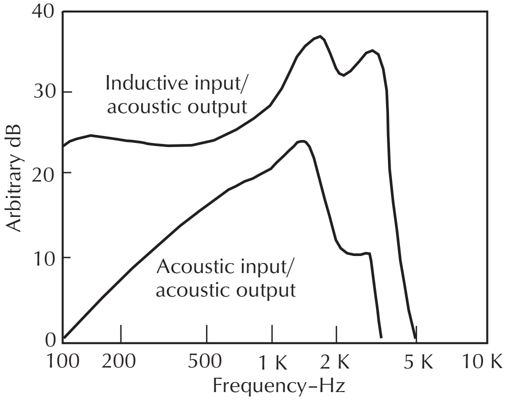


Figure 42-1. Typical hearing aid response. (From Reference 1.)

The field strength should be as uniform as possible over the coverage area. An achievable criterion for uniformity is a maximum variation of ±3 dB in the audio output signal.

System design is based on the vertical component of the magnetic field, ignoring the horizontal field components for three reasons:

- The vertical field strength predominates over most of the loop area, Fig. 42-2.

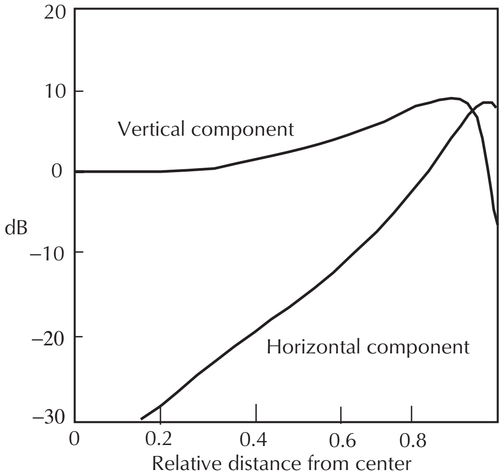


Figure 42-2. Field strength along the diagonal of a square loop. (From Reference 2.)

- The T-coils in hearing aids are typically positioned to be most sensitive to the vertical field.
- Rotating the hearing aid about the vertical axis (as in turning the head) results in no change in the pickup of the vertical component, whereas the pickup of the horizontal component changes from zero to maximum to zero with such rotation.

42.2.2 Loop Location and Size

The field strength produced by the loop will vary in intensity from the edge of the loop to the center, Fig. 42-3. The range of variation is dependent on the area and shape of the loop and the listening height, which is the vertical distance between the plane of the loop and the receiver. This interrelationship is expressed as the relative listening height and is determined by

$$h_r = \frac{h}{\sqrt{0.5A}} \tag{42-1}$$

where,

h_r is the relative listening height,

h is the listening height,

A is the area covered.

The normalized field strength along the diagonal of loops of various shapes and the corresponding range of acceptable values for h_r are shown in Fig. 42-4.

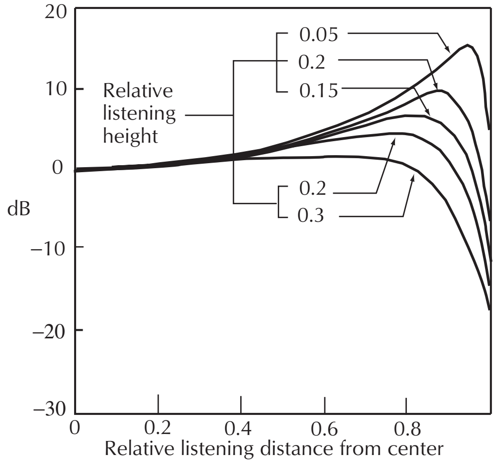


Figure 42-3. Vertical field strength along the diagonal of a square loop. (From Reference 2.)

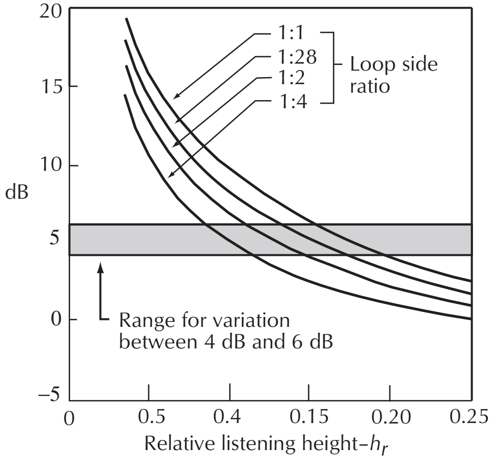


Figure 42-4. Vertical field strength along the diagonal of rectangular loops. (From Reference 2.)

By application of Eq. 42-1 and the h_r values found in Fig. 42-4, it is possible to design a loop of acceptable shape, area, and listening height. The penalty for an inequality in Eq. 42-1 is degraded uniformity of field strength, as can be seen in Fig. 42-4.

If the loop is to be placed at floor level (h = 48 inches for seated listeners), square loops falling within the acceptable h_r range will vary from 28 ft × 28 ft to 38 ft × 38 ft. A rectangular 1:4 loop may range in size from 24 ft × 96 ft up to 32 ft × 126 ft. Smaller loop dimensions will require a smaller h; larger areas need a larger h.

As h grows larger, the field-distorting effect of steel in the building structure and in the audience chairs becomes more pronounced. This field distortion is manifest by dead spots within the loop. At its worst, the entire system may be rendered useless.

The great listening height required of large loops also presents architectural problems. Often the only practical place to locate a loop is at floor level, either below the floor or under the carpet. Where it is not feasible to locate the loop far above (or far below) floor level, the single, large loop can be broken up into a number of smaller loops that can be sized to locate at floor level. Because the vertical field strength rapidly falls to a minimum above the conductors, it is important to locate the loop wires in aisles or other areas that do not require coverage. For multiple loops, the current in parallel conductors of adjacent loops must flow in the same direction, Fig. 42-5.

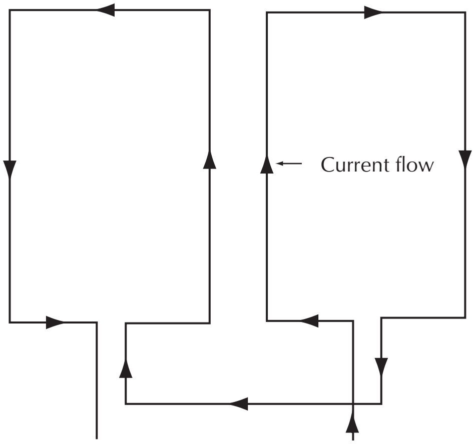


Figure 42-5. Multiple-loop current flow diagram.

Unfortunately, multiple loops will almost always have poorer uniformity than a single loop of the same size as one of the multiples. There is a special design technique for achieving a more nearly constant vertical field strength when using multiple loops. It involves the use of two sets of overlapping loops that are driven with electrical signals 90 degrees out of phase. This complex procedure is described by Bosman and Joosten.[3]

42.2.2.1 Loop Current

Once the size and location of the loop are fixed, the required current in the loop can be calculated. The strength of the magnetic field is directly dependent on the current in the loop. The required current, *I,* in a single-turn loop is

$$I = \frac{0.1\,\text{A/m}^{*}\pi A}{2D} \times (1 + 2h_r^{\,2}) \times \sqrt{1 + h_r^{\,2}} \qquad (42\text{-}2)$$

* 0.0305 A/ft in English units

where,

0.1 A/m is the field strength criterion,

A is the loop area in square meters or square feet,

D is the loop diagonal in meters or feet.

The terms containing h_r are a correction for the distance of the listener from the plane of the loop and are obtained from Fig. 42-6 by going vertically from h to the line and horizontally to the correction distance.

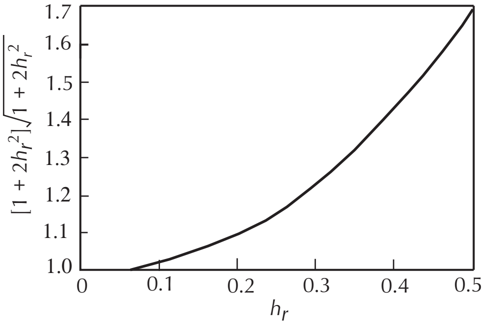


Figure 42-6. Graph for obtaining distance correction.

If a multiturn loop is used, the required current in the loop is

$$I_M = \frac{I}{n} \qquad (42\text{-}3)$$

where,

I is the current from Eq. 42-2,

n is the number of turns.

42.2.2.2 Loop Impedance

Wire size and number of turns in the loop must be selected to handle the required current safely and to control the range of variation of impedance across the audio band. A loop can be designed to provide the required magnetic field strength by using a relatively small wire with one turn or by using a larger wire with several turns. In the first case the loop impedance would be mainly resistive; in the second, it would be heavily inductive.

The impedance increases with frequency because of the inductive reactance of the loop. This increase is limited by adjusting wire size and the number of turns so the impedance at 1000 Hz is no more than three times the impedance at 100 Hz. This moderately rising impedance characteristic and falling loop current will complement the rising sensitivity characteristic of the T-coil, Fig. 33-7. Too high an impedance at high frequencies will result in too low current, producing poor response and degraded SNR.

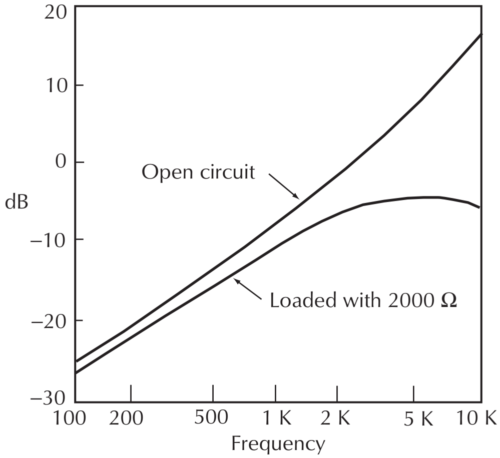


Figure 42-7. Sensitivity of typical inductive coil. (From Reference 1.)

A wire size that can handle the required current with an acceptable heat rise is selected from Table 42-1. The impedance is then calculated at several frequencies such as 100 Hz, 1 kHz, and 10 kHz. The following equations are useful:

$$L = \frac{rn^2}{13.5}\log\frac{2.8r}{d} \tag{42-4}$$

where,
L is the inductance in henrys,
r is the radius of the loop in inches,
n is the number of turns,
d is the diameter of the conductor in inches
(this is a simplification of Wheeler's equation);

and

$$Z = \sqrt{R^2 + (2\pi f L \times 10^{-6})^2} \tag{42-5}$$

where,
Z is the loop impedance,
R is the dc resistance of the length of the coil,
f is the frequency of interest,
L is the inductance of the loop from Eq. 42-4.

An example of the calculations required to determine the impedance of a one turn 20 ft × 20 ft loop of AWG #20 wire at 1 kHz is

$$\begin{aligned} L &= \frac{(10\text{ ft} \times 12\text{ ft}) \times 1^2}{13.5}\log\left[\frac{2.8(10\text{ ft} \times 12\text{ inches})}{0.03196}\right] \\ &= 143\text{ H} \end{aligned}$$

$$\begin{aligned} Z &= \sqrt{\left(80 \times \frac{10.15}{1000}\right)^2 + (2\pi \times 10^3 \times 143 \times 10^{-6})^2} \\ &= 1.21\ \Omega \end{aligned}$$

If this is connected to a 1:4 autotransformer, the amplifier will see 4.84 Ω

Throughout this design procedure there is a certain amount of approximation involved; for instance, Eq. 42-4 applies to round loops. Error is introduced in calculating the inductance of a square or rectangular loop, however, this error is not great enough to seriously affect the results.

42.2.2.3 Electronic System

The power amplifier is selected that can supply the required current to the loop. The power required is determined with the basic equation

$$P = I^2 Z \tag{42-6}$$

The adjustment of the output current is determined by the equation

$$I = \frac{E}{Z} \tag{42-7}$$

An autotransformer or other suitable impedance matching device is used to match the amplifier to the loop. In the absence of an impedance meter, the loop dc resistance may be used for matching to the amplifier because at the minimum impedance point (low frequencies), the impedance is largely resistive.

A typical loop system diagram is shown in Fig. 42-8. In very large halls, a delay unit may be required in the more distant loops in order to avoid excessive time delays between the loop signal and the acoustic signal. Equalization is desirable to compensate for any frequency response irregularities. The equalizer is adjusted to provide a natural sound quality with a typical receiver and to insure that power is not transmitted outside the power bandwidth of the receiver.

A compressor is needed to insure that the system does not produce excessive distortion at high signal levels, either from clipping the amplifier or from overloading the hearing aid T-coil. The compressor should

Table 42-1. Copper Wire Data

AWG #	Ohms/1000 ft	Diameter in Inches	Current for Heat Rise*			Melting Current
			5°C	10°C	20°C	
24	25.67	0.02010	2.3 A	3.2 A	4.5 A	29.2 A
22	16.14	0.02535	3.0 A	4.2 A	5.9 A	41.2 A
20	10.15	0.03196	4.0 A	5.5 A	7.8 A	58.6 A
18	6.375	0.04030	5.5 A	7.8 A	10.0 A	82.4 A
16	4.016	0.05082	7.5 A	10.0 A	15.0 A	117.0 A

*Heat rise based on an insulation thickness of 10 mils. Heavier insulation allows more current for the same heat rise.

be adjusted according to the nature of the principal program material. If the system is used mostly for music, a compression ratio of about 4:1 will result in minimal harm to the music. If speech is the principal program, compression ratios up to 20:1 can be used to improve both intelligibility and SNR.

42.2.2.4 Installation

If the loop is installed in conduit, it must be nonmetallic conduit such as PVC and should be placed so that there is little (or no) steel between the loop and the listener. Often the conduit is run in the top of a concrete slab or below a wood-framed floor; but it can also be run in walls or even the ceiling of a room. When installing a loop in an existing room, it is often easiest to run the loop wire under the carpet, using conduit only for the run to the amplifier.

42.2.3 FM Broadcast

FM broadcast systems have replaced many magnetic loops in classrooms where hearing-impaired children are taught because the FM signal is normally free from noise and provides a more uniform and reliable signal. Several channels are available so systems can be used in adjacent rooms. The sound quality is excellent. The useful receiving range will vary from 30–90 m (100–300 ft) depending on the amount of steel in the building. Transmitters are available for operation from the powerline for permanent installations or by battery for portable applications.

The Federal Communications Commission (FCC) has set aside a band of frequencies, 72.025–75.975 MHz, for FM broadcasting to the hearing impaired under FCC Rules Part 15. These frequencies cannot be used for any other purpose, such as language translation systems or communications applications. No license is required, although the manufacturer of the transmitter is required to have FCC approval of the transmitter design. The FCC restricts radiation to a maximum field strength of 8000 µV/m at 30 m. The FCC rules require a special antenna connector on the hearing assistance transmitters to prevent the use of illegal gain antennas that could result in a higher transmitted field strength than dictated by the FCC. The system requires no special knowledge to install; sufficient instructions are provided by the manufacturers.

The FCC has opened the 216–217 MHz band for assistive listening devices. This band falls under the *low power radio services* (LPRS) of the FCC, which limits the power output to 100 mW. While the 72 MHz band could transmit 500–1500 ft, the 216 MHz band can transmit 1000–3000 ft.

The systems can be either wide-band or narrow-band. Wide-band systems have the following characteristics:

- High fidelity for all applications.
- Low cost.
- Good rejection of unwanted or external radio signals.
- Limited to six simultaneous channels.

Narrow-band systems have:

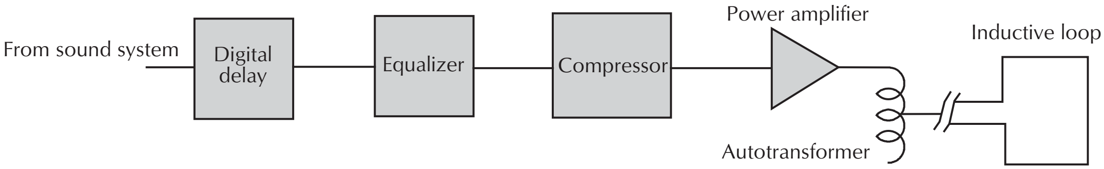


Figure 42-8. Induction loop system block diagram.

- High immunity to unwanted or external radio signals.
- More than ten simultaneous channels.
- Good fidelity for voice applications.

FM broadcast is much like wireless microphones; more information can be found in Chapter 16.

An FM assistive listening system by Listen Technologies LS-04 Installed System includes programmable receivers, and a charging case and rechargeable batteries, Fig. 42-9. The LR-500 receiver can be programmed to receive only the channels available at the venue. This system helps public venues to meet Americans with Disabilities Act (ADA) guidelines. It is easy to add receivers and a wireless speaker/receiver.

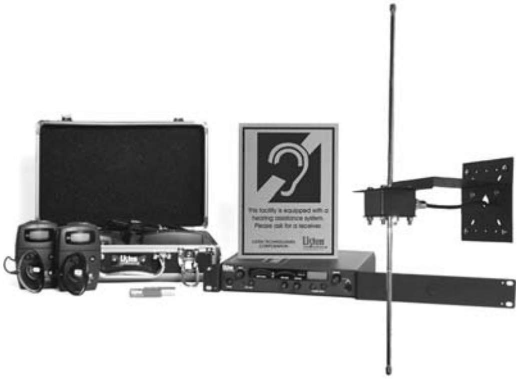
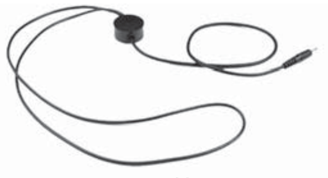


Figure 42-9. An installed FM assistive listening system. Courtesy Listen Technologies Corporation.

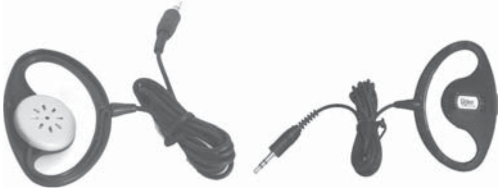
The system has an SNR of 80 dB and is available for 72 MHz, 216 MHz, or 863 MHz band. The system includes one LT-800 stationary transmitter, an antenna kit, rack mount kit, and four LR-500 programmable display receivers with ear speakers.

The Personal System by Listen Technologies includes a portable transmitter and receiver in a soft-sided carrying case. The Personal System's LT-700 portable transmitter, lapel mic, LR-400 display receiver, and ear speaker all fit in a soft case so they can be taken to school, house of worship, or theater. The listener gives the transmitter with its microphone to the presenter and the listener uses the receiver with an LA-166 neckloop or LA-164 earphone. The neckloop generates a magnetic field that is picked up by hearing aids that are equipped with a T-coil, Fig. 42-10.

The Sennheiser Mikroport 2015 is suited for classroom use and allows a hearing impaired student to have an improved learning experience via wireless audio connection to the teacher. The system includes a wireless transmitter with a lavalier microphone worn by the teacher and a body worn wireless receiver for the student. Direct audio input cables are available for use with cochlear implants and hearing aids and induction neck loops for use with T-coil hearing aids. The systems can also be used with standard headphones or ear buds. Multiple receivers can be used with a single transmitter and, because there are hundreds of discrete frequencies available, systems can be used in adjoining classrooms without crosstalk or interference, Fig. 42-42-11.



A. Neckloop.



B. Earphone.

Figure 42-10. Neckloop and earphone. Courtesy Listen Technologies Corporation.

42.2.4 AM Broadcast

AM broadcast systems are largely unregulated by the FCC in the United States. The International Standard IEC118 Parts 1 and 4 define the performance, field strength, frequency response, and spurious levels for induction loop and hearing aids. The European Standard EN 60118 is also applicable in Europe. The basic rules (from FCC Bulletin OEC 12 dated July 1977) are:

- The system must not cause any interference to an existing licensed service.
- The operating frequency must be in the AM broadcast band or below (10 Hz–490 kHz and 510–1600 kHz).
- An open-wire antenna may be used in the 510–1600 kHz band providing it does not exceed 10 ft in length and providing the transmitter is restricted to an input power of 100 mW. Any type of antenna and transmitter power may be used provided the field strength does not exceed:
  - ($2400/F$) µV/m at 300 m in the 10–490 kHz band.
  - ($24{,}000/F$) µV/m at 30 m in the 510–1600 kHz band where F is the carrier frequency in kHz.

Figure 42-11. Personal system including FM transmitter, receiver, and battery charger. Courtesy Sennheiser Electronic Corporation.

- Carrier current system radiation limits are set differently. The radiation from the electrical system must not exceed 15 µV/m at a distance of 157,000/*F* ft, where *F* is the frequency in kHz.

An examination of these radiation restrictions indicates that the lowest operating frequency results in the greatest coverage area. AM systems typically operate on carrier frequencies below 700 kHz.

The sound quality and noise level of an AM broadcast signal are better than from a magnetic induction loop but inferior to an FM system. In general, if a pocket AM radio receives a local station well in the space, the low-power broadcasting system will work as well. Systems operating in the commercial broadcasting band (540–1600 kHz) may be picked up on any AM receiver. Systems operating below the standard broadcast band are available with fixed-frequency, nontunable receivers.

At first it may seem economically attractive to select a system operating in the regular broadcast band because the patrons could be expected to furnish their own receivers. But on reflection, there is much to be said in favor of the special nontunable receiver as the patron is not required to find the signal among the many public radio stations.

The coverage area from an AM system depends on the type of antenna employed. The three types of antenna are *open wire*, *lossy coaxial cable*, and *carrier current*.

Open-wire antenna systems, with their restricted power and antenna length, usually achieve coverage ranging from 15–45 m (50–150 ft) depending on the amount of steel in the building. Fixed-frequency, below-the-broadcast-band systems usually employ an open-wire antenna. These systems are used by many churches.

Lossy coaxial cable systems employ a special type of coaxial cable that has a very loose or open shield, allowing a little radiation to occur all along the cable. The usable reception range is 15–23 m (50–75 ft) from the cable. The length of the cable may be quite long. The lossy coaxial technique is found in drive-in cinemas, stadiums, and arenas and can be used for broadcasting flight arrival and departure information along the highway approaches to airports.

The most common type of AM system is carrier current. In this system the output of the transmitter is capacitive coupled into the main power distribution wiring of the building. The radio signal travels throughout the building on the electrical wiring. This system is widely used on college campuses for limited-coverage, student-operated radio stations. Carrier current is an inexpensive way to provide program monitoring throughout a building.

The costs of the low-frequency, open-wire transmitter and fixed-frequency receivers are about the same as FM systems. The lossy coaxial and carrier-current systems may cost more, depending on the power required and the length of the coaxial cable. Any of these systems can be installed easily by following the instructions from the manufacturer. Both the lossy coaxial and carrier-current systems should be planned through consultation with the manufacturer of the transmitter.

42.2.5 Infrared

Infrared light can be used to broadcast a very high-quality signal. Presently available systems can broadcast up to twelve different programs on the same emitter, making infrared very useful for large-scale language translation systems. Infrared systems are also used in museums and for lecturing and teaching on auscultation, the listening of heart beats. Systems are also produced for home listening, for both stereo and for video (TV). It is the only system for the hearing impaired that can transmit in stereo. Also unlike the other systems, infrared broadcasts are completely contained within the room because infrared behaves like visible light; it cannot go through a wall; even a heavy cloth is an opaque barrier. This control of the broadcast range is a significant factor where confidentiality is important, as in corporate meeting rooms. For this reason and because of its outstanding sound quality, infrared is the system of choice for professional theaters and concert halls.

42.2.5.1 Components of an Infrared System

The basic components of an infrared system are:

- The control transmitter (which is connected to the audio source)
- Slave emitters, daisy-chained together (if needed).
- The receivers.

42.2.5.2 Categories

Any installation generally falls into four categories:

1. Personal listening systems (PLS) for living rooms, bedrooms, offices, etc.
2. Medium area systems (MAS) for lounges, meeting rooms, courtrooms, classrooms, small theater, churches, etc.
3. Large area systems (LAS) for auditoriums, large theaters, churches, arenas, etc.
4. Large area multichannel Systems (up to 32 channels) for simultaneous interpretation and other applications.

42.2.5.3 Coverage

Transmitter and emitter panels with present state-of-the-art design and components allow over 70 ft^2 of coverage per IR light emitting diode in single channel IR systems. The shape of the polar pattern of a panel is nearly identical to the pattern of a single diode. The half-power angle of luminosity for the presently used LEDs is approximately ±25 degrees. More LEDs increase the IR intensity, and the area of coverage multiplies by the number of diodes in the panel. The radiation pattern can be considered the same for horizontal and vertical orientations, and is scarcely influenced by the arrangement of diodes or the housing of the array.

Since there is a physical limit to the light output power of any LED, the total output has to be shared between channels in a multichannel system and the available coverage area per radiator has to be divided by the number of channels in the system. Conversely, for the same required coverage, the number of radiators should be multiplied by the number of channels in the system, or twice the amount of emitters in a stereo system as would be required in a single channel installation in the same venue.

Reflection and scatter off walls, ceilings, floors and furnishings broaden the coverage and make it largely nondirectional. Infrared light behaves a lot like visible light as it reflects best off bright and smooth surfaces like white walls, and is absorbed by dark and rough materials like black velvet curtains.

Emitters should be placed in a manner to provide even illumination throughout the room. The are usually mounted 10–40 ft above the floor and pointed down toward the audience. When emitters are placed on both sides of the stage, they should be cross-fired into the audience. Any number of receivers can be used in the system as they will not effect the signal source.

LEDs have degradation of light output over time, however, by using good electronic circuit design, a projected continuous operating life under standard conditions of more than 100,000 h, before the light output diminishes to 70% of its original value, can be obtained.

42.2.5.4 Ambient Light

Infrared systems work in virtually any environment except for direct sunlight. Systems can even be installed in shaded outdoor areas. Rooms with very high ambient light levels or poorly filtered fluorescent ballasts may require additional emitters for a sufficient SNR.

42.2.5.5 The infrared Link

Infrared systems can be either narrow band or wide band, depending on your requirements. Table 42-2 gives the technical specifications for infrared systems.

The infrared link uses a specially doped gallium arsenide light emitting diode (LED) to transmit the signal. Each diode emits about 10 mW total radiant power, requiring up to 143 LEDs in each array to produce adequate power. The wavelength of the emitted light is 930 nm and is neither monochromatic nor coherent, so any number of diodes can be used together without interference between them. The useful coverage pattern of the emitter varies with distance and the number of channels being transmitted, Fig. 42-12. The number of emitters required depends upon the size and shape of the area to be covered and the number of channels in use. Emitters are usually employed in pairs, located at each front side of the audience and cross-fired across the seating area so that each person receives an infrared beam from each side. This cross-firing helps to eliminate shadowing from other people in the audience.

Table 42-2. Technical Specifications for Infrared Systems

Characteristic	Narrow Band	Wide Band
No. of channels	12	2
Carrier frequencies	55–535 kHz, excluding 455 kHz	95 kHz and 250 kHz
Channel spacing	40 kHz	155 kHz
Modulation	FM	FM
Pre-emphasis	100 µs	50 µs
Normal deviation	±6 kHz	±35 kHz
Peak deviation	±7 kHz	±50 kHz
Transmitters		
Frequency response (–3 dB)	50 Hz–8 kHz	50 Hz–13 kHz
Max. distortion (1 kHz)	<1.0%	<1.0%
SNR (A weighted)	>55 dB	>70 dB
Receivers		
Frequency response (–3 dB)	50 Hz–8 kHz	100 Hz–9 kHz
Max. distortion 1 kHz)	<2.5%	<1%
SNR (A-weighted)	>55 dB	>63 dB
Emitter Panels		
Frequency response (–3 dB)	30 Hz–710 kHz	30 Hz–710 kHz

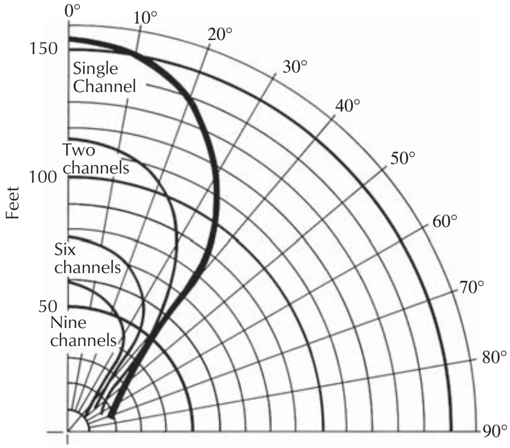


Figure 42-12. Infrared emitter panel coverage patterns.

The receiving end of the infrared link is a silicon photo diode that is reversed biased and produces current when struck by photons. The light gathering area is small, 7 mm^2, but is effectively increased by mounting it in a collecting lens, Fig. 42-13.

The silicon PIN receiving diode has a maximum sensitivity at a wavelength of 850 nm. Fig. 42-14 shows the spectral sensitivity of the eye, IR LED transmitting diode, IR filter, and receiving diode.

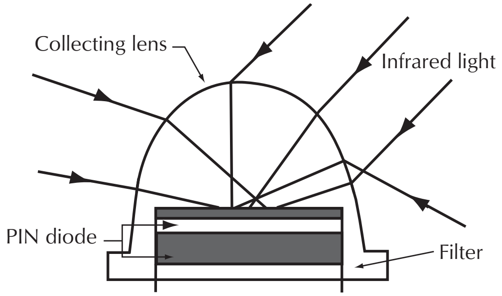


Figure 42-13. Infrared receiving diode.

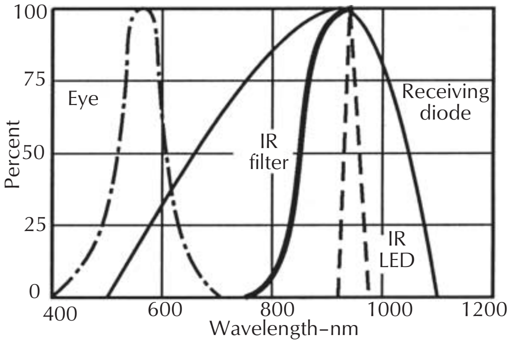


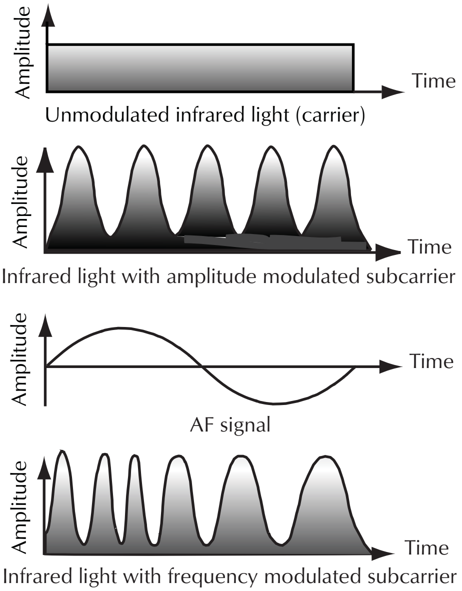
Figure 42-14. Sensitivity of the eye, IR LED transmitting diode, IR filter, and receiving diode.

Infrared light behaves much like visible light; it reflects off of light-colored walls, ceilings, and other surfaces so a receiver can "see" the signal even without direct line-of-sight to the emitter. Also, the receiver's ultrawide-angle fisheye lens captures direct or reflected signals from almost any direction.

There is a unique limitation to the use of infrared: it cannot be used in bright daylight. The infrared light occurring as a natural part of daylight will override the lower power-modulated light from the system. The system can also receive interference from very high-level incandescent lamps. Partially dimmed incandescent lamps can also be a problem in some situations because the reduced voltage to the lamps causes a shift toward red that greatly increases the infrared output of the lamp. This increase in infrared interference can, on rare occasions, be a problem where audience down lights are left at a dimmed setting and the infrared beam from the system is weak. Deep under a balcony is a likely trouble spot. When this problem occurs, it is necessary to dim the lights more or add more emitters to the infrared system to cover under the balcony.

An infrared system comprises three sections: the transmitter, the emitter (sometimes both combined in one unit), and the receiver. The transmitter imparts the audio signal onto a subcarrier signal which the emitter converts into infrared light. The receiver decodes the infrared signal to retrieve the original audio.

To achieve a usable radiated power level, the IR LEDs are used in multiple arrays. Their light output is amplitude-modulated by one or more frequency-modulated subcarriers (typically 95 kHz for single-channel wideband systems; 95 kHz and 250 kHz for two-channel systems). Each channel's audio signal frequency-modulates its particular subcarrier, Fig. 42-15.



Unmodulated infrared light (carrier)

Infrared light with amplitude modulated subcarrier

AF signal

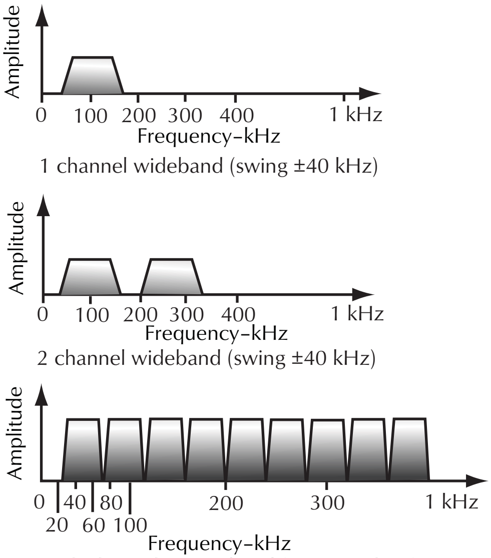
Infrared light with frequency modulated subcarrier

Figure 42-15. Infrared system's modulation technique.

Two transmission modes are available: wideband, for one or two channels of high-fidelity audio; or narrowband, for up to twelve channels with a 70–7000 Hz response suitable for communications, Fig. 42-16.

Since the transmission medium is a modulated carrier of harmless invisible light, Fig. 42-17, instead of radio or audio signals, it is immune to outside interference and also causes none itself. No operator licensing is required for use of infrared systems.

Manufacturers provide detailed instructions for planning and installing the system. Typical installations are shown in Fig. 42-18. The rear emitters must cover both under the balcony plus the balcony. For this reason, separate emitters may be required to cover both areas. The advantages of infrared systems are fully realized in applications where different audio programs are required in adjacent rooms, such as a multicinema complex. Each room can be equipped with the same system without interference among them. No frequency coordination is required as with radio frequency systems. The same receiver can be used in any theater.



1 channel wideband (swing ±40 kHz)

2 channel wideband (swing ±40 kHz)

Multichannel narrowband (swing ±7 kHz)

Figure 42-16. Infrared systems channel allocation.

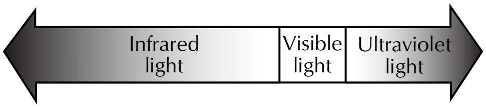


Figure 42-17. Light spectrum.

A useful tool in aiming the emitters is a low-cost, black-and-white television camera and monitor. Most monochrome television cameras have useful sensitivity in the infrared region. With the room lights off, observe on the television monitor the part of the room illuminated by the infrared beam. The well-illuminated area will be the area of good reception. A corollary to this procedure is that the infrared television viewing system can be used to view a darkened stage—for instance, for coordination of rigging and prop moves in a fast, complicated change in the dark.

42.3 Receivers

Receivers are required with all systems, though fewer are needed with an induction loop because many patrons will have aids equipped with T-coils. Most manufacturers of systems for the hearing impaired offer several types of earphones with their receivers. Typically these include a single earpiece, a stethoscope-type dual earpiece, Fig. 42-19, and an induction loop for use

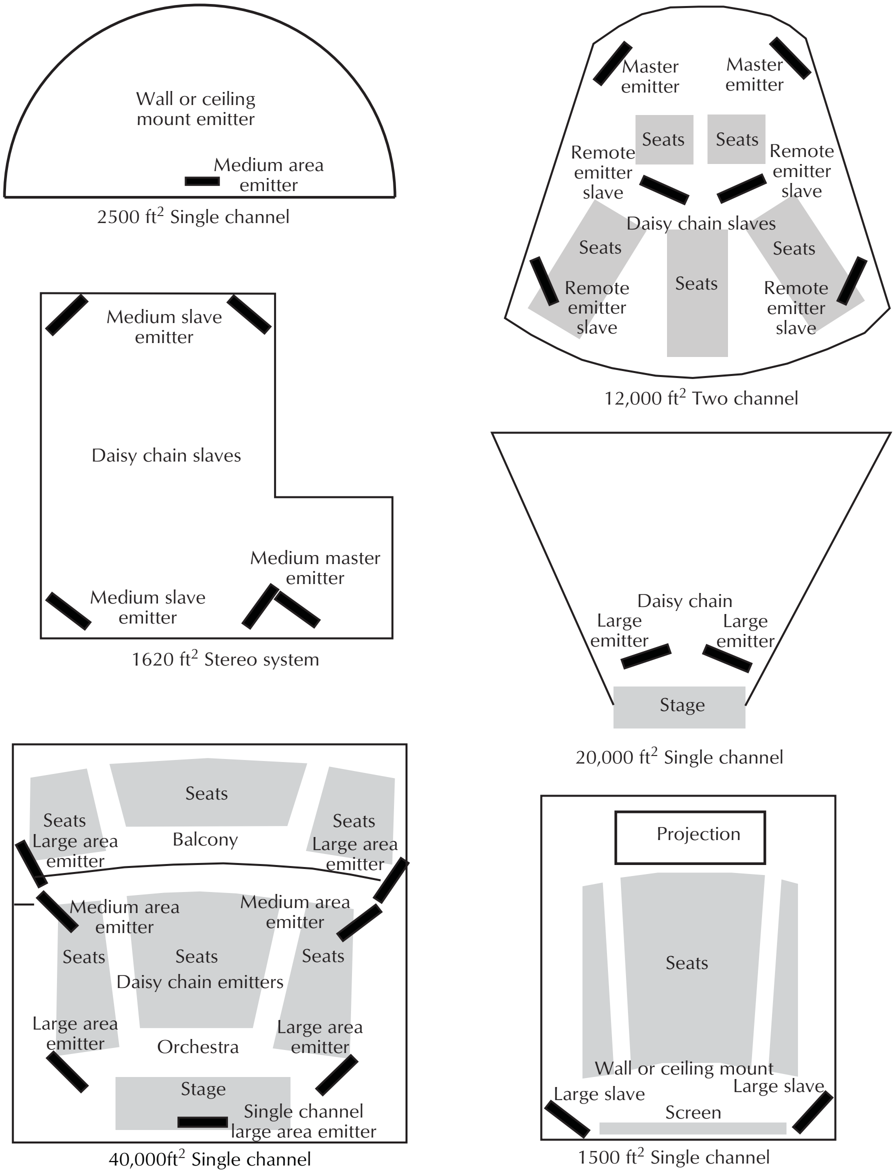

Wall or ceiling mount emitter
Medium area emitter
2500 ft² Single channel
Master emitter
Master emitter
Seats
Seats
Remote emitter slave
Remote emitter slave
Daisy chain slaves
Seats
Seats
Seats
Remote emitter slave
Remote emitter slave
12,000 ft² Two channel
Medium slave emitter
Daisy chain slaves
Medium master emitter
Medium slave emitter
1620 ft² Stereo system
Daisy chain
Large emitter
Large emitter
Stage
20,000 ft² Single channel
Seats
Seats
Large area emitter
Balcony
Seats
Large area emitter
Medium area emitter
Medium area emitter
Seats
Seats
Seats
Daisy chain emitters
Large area emitter
Orchestra
Large area emitter
Stage
Single channel large area emitter
40,000ft² Single channel
Projection
Seats
Wall or ceiling mount
Large slave
Large slave
Screen
1500 ft² Single channel

Figure 42-18. Typical infrared system theater installation.

with the patron's T-coil. Most users report a strong preference for the dual earpiece instruments and also report a strong dislike for an over-the-head-type headphone because they are uncomfortable for long periods of use and destructive to hair styles. Two types of induction loops are available. One is a small coil that hooks over the ear close by the hearing aid; this type is often hard for elderly people to use properly. The more popular loop is a lanyard type that hangs around the neck and may be used to support the receiver.

Many theaters and churches that have installed a system for the hearing impaired have found normal-hearing patrons using the system to enhance their listening comfort, which is especially true in larger houses with seats in areas of poor natural acoustics. These normal-hearing patrons universally prefer the dual earpiece both because it sounds more natural and because a single earpiece leaves one ear open to receive the live sound from the stage with a signal-delay annoyance. Depending upon the distance from the stage, this delay can be very distracting to those with less than profound hearing impairments.

Battery replacement and earpiece sanitizing are the principal maintenance problems with all systems for the assisted listening devices. Batteries may last up to one year, although that seems an uncommonly long life if the receivers are being used often. The infrared receivers have rechargeable batteries, which should be recharged after each use. Earpieces are most commonly sanitized by replacing the plastic ear tips or by using replaceable foam balls.

Most theaters and concert halls provide receivers to their patrons for no charge or for a small fee to cover the cost of handling, batteries, and sanitizing. Some organizations have been successful in selling receivers to regular patrons, especially in communities where several theaters and churches use the same technology.

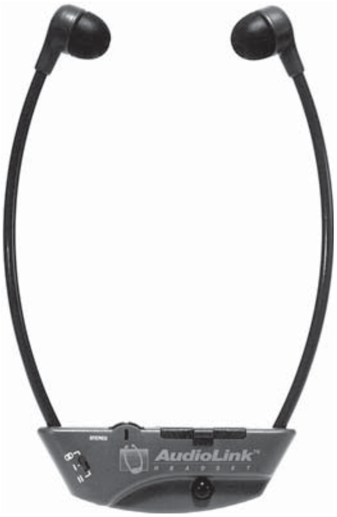


Figure 42-19. A wireless IR system receiver under the head multichannel headset. Courtesy Sennheiser Electronic Corporation.

References

1. M. C. Martin, “Power Requirements for Inductive Loops,” *Royal National Institute for the Deaf*, London, England.
2. S. C. Dalsgaard, “Field Distribution Inside Rectangular Induction Loops,” State Hearing Centers, Research Laboratory for Technical Audiology, University Hospital, Odente, Denmark.
3. D. Bosman and L. J. M. Joosten, “A New Approach to a Space-Confined Magnetic Loop Induction System.” *IEEE Transactions on Audio,* May/June 1965.

Bibliography

K. Borrild, “The Induction Loop and Its Possibilities,” The State Boarding School for Hard of Hearing and Deaf, Frederica, Denmark.

E. deBoer, C. Kruidenier, and K. W. Kortschot, “On Communication Systems in Schools for the Hard-of-Hearing,” *Acustica*, vol. 19, no. 3, 1967–1968.

RNID Scientific and Technical Department, Installing a Loop System. “Hearing,” *Royal National Institute for the Deaf,* October 1977.

This page intentionally left blank

Chapter 43

Intercoms

by Glen Ballou

43.1 General Purpose Intercoms 1561
  43.1.1 Point-to-Point Systems 1561
  43.1.2 Matrix Systems 1562
  43.1.3 Apartment Security/Intercom Systems 1563
  43.1.4 Residential Intercom Systems 1564
  43.1.5 Commercial Security Systems. 1564
    43.1.5.1 Zone Paging 1564
    43.1.5.2 Security Audio Monitoring 1564
    43.1.5.3 Emergency Crisis Communications 1566
  43.1.6 Commercial/Industrial Systems 1566
  43.1.7 Wireless Intercoms 1569
43.2 Broadcast Intercoms 1569
  43.2.1 Broadcast Point-to-Point Systems 1570
  43.2.2 Broadcast Party Line System 1570
  43.2.3 Broadcast Matrix Systems 1571
  43.2.4 Program Interrupt (IFB) 1573
  43.2.5 Wireless Broadcast Intercom System 1574
  43.2.6 Belt Packs 1575
  43.2.7 Telephone Interface 1575
  43.2.8 Headsets. 1575

This page intentionally left blank

Intercoms, short for intercommunication systems, are often more sophisticated than sound reinforcement systems because sound systems are unidirectional while intercoms are bidirectional. Sophistication varies from the simple home intercom to hotel and apartment intercoms, to hospital nurse call and emergency systems, to party line systems, to the multiprocessor controlled, multimemory, analog and digital intercommunication systems.

43.1 General Purpose Intercoms

Intercom systems fall into three categories: point-to-point, party line, and matrix, Fig. 43-1. The *point-to-point* or *dedicated line* is a private line between two stations. Normally, no other stations can hear your conversation.

The *party line*, *conference line,* or *distributed line* is a shared line where a number of people, usually sharing a common task, are all talking to each other. This is like being at a conference table where everyone can talk and listen to one another at the same time and there is no privacy capability. By having multiple circuits and multiple wires, this system can communicate to different parties at the same time. Most often this type is found in broadcast intercoms.

Most major intercom manufacturing companies are offering digital *matrix* systems. These systems are multiprocessor controlled, multimemory, and have analog and digital audio and can combine the point-to-point and party line communications together. All have advantages and disadvantages and they can be either hard-wired or wireless or a combination of the two.

Intercom systems can be simple, a two unit system with only a call button and a microphone/loudspeaker combination in each unit, to a complicated multichannel multistation system with a separate microphone and loudspeaker, display window, keypad, TV entrance monitor, auxiliary inputs, can be programmable, plus have a multitude of special features.

43.1.1 Point-to-Point Systems

Point-to-point systems are the simplest type of intercom and are mostly used in residential and office applications, schools, apartments, and nurse call and emergency call systems. With point-to-point, a caller or originator makes contact with the desired receiver or receivers and communicates only with them, Fig. 43-1. All other stations are isolated so they cannot be part of the one- or two-way conversation. This is like the telephone system, you dial a party or parties through a conference call, and the conversation can only be heard by those parties that were called up even though the telephone system has hundreds of thousands of other phones on the system.

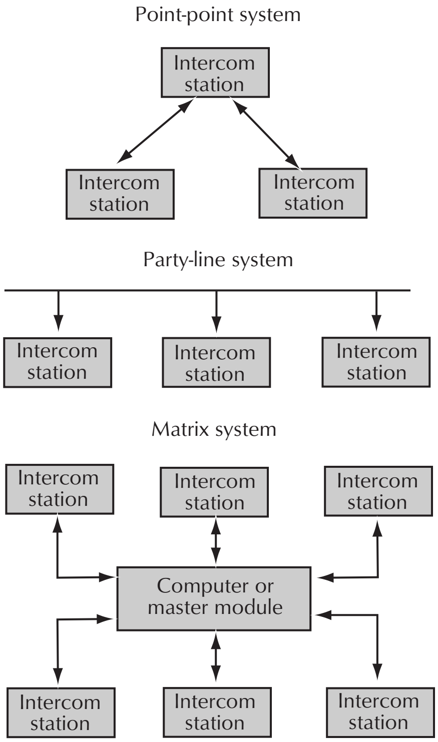
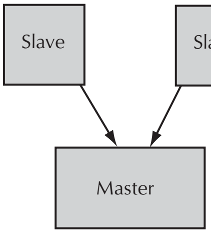


Figure 43-1. The three basic types of intercoms.

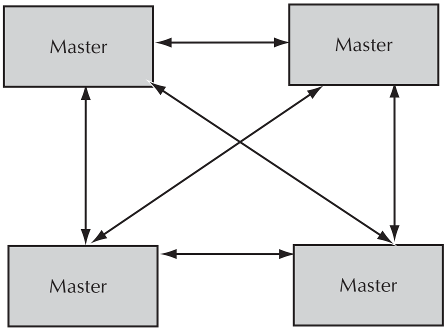
Systems are normally made up of master stations, slave stations, door monitoring and opening stations, and input devices such as AM/FM receivers.

These systems have a variety of features and are used for intercommunications between rooms or desks, intercommunications between an indoor station and an outside door including door release, and provide some form of playback signal at all or selected stations in the system. The system can be one master to one or more slaves, all master, or a combination of the two. Masters have the ability to originate calls to any or all stations while slaves can only call the master station they are connected to, Fig. 43-2. Stations can be in-the-wall units or selfcontained desktop units, hardwired or wireless. Normally hardwired systems are the best for point-to-point systems as they are less expensive and less likely to have interference. The disadvantage is they are more difficult to move or reposition.

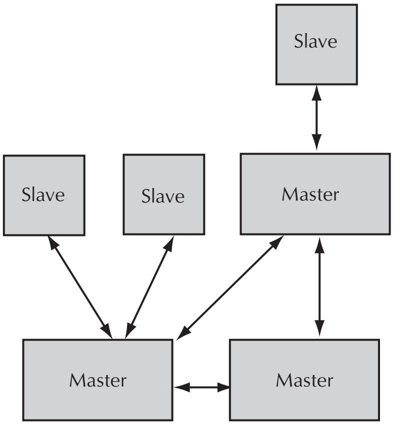
The Bogen Model PI35A, Fig. 43-3 is a twenty five-station intercom. It provides facilities with two-way communication, emergency paging, time tones, and



Master/slave



Master/master



Intermixed

Figure 43-2. Master/slave, master/master, and intermixed intercoms.

background music or other audio program material to speaker-equipped locations. The Bogen PI35A is designed to meet highpower paging and intercom requirements with features suitable for applications with mixed noise environments (construction, retail stores, small factories, parking garages, etc.). A 20 W intercom amplifier features a voice-shaped frequency response for intelligibility. A 35 W program amplifier is used for program material and/or emergency announcements.

Figure 43-3. Bogen P135A twenty five station intercom. Courtesy Bogen Communications, Inc.

Program material from microphones, a CD player, or other background music source can be used. Distribution is accomplished by simple push-button selection. Emergency announcements take precedence over program distribution and are accomplished with a single push-button selection. A time signal can also be sent to all stations. Optional paging facilities permit emergency all-call paging from a remote telephone, or interphone or microphone. Telephone paging captures system priority and overrides all system functions except the emergency page feature.

A room selector panel is provided to select intercom and program functions for each station. Calls from stations are initiated through call-in switches in the various rooms and are announced at the control center by light and tone annunciation. The system provides a 25 V balanced output and operates from a 120 Vac, 60 Hz source. The system consists of a master control panel and a twenty five-station room selector pane. A number of options are available including room call-in switches, call-in adapter modules for existing older call-in switches, two-wire call-in adapters, and various styles of transformer-coupled loudspeakers.

43.1.2 Matrix Systems

Not too many years ago, all intercoms used a mechanical switching matrix to call and to route the voice signal. Systems were simple, wires came from a central system and went out to the individual masters and submasters. Switches were multipole and carried signal and voice. Often the voice line was shielded to eliminate hum and stray noise pickup, however, on inexpensive systems unshielded lines were used and the frequency response of the receiver was limited to 200 Hz–4 kHz to eliminate noise and hum.

While these systems are still used today, many have been replaced with digital matrixes and electronic switching. These systems use low voltage/low power and because the signal is digital, seldom require shielded lines.

Matrix systems can control upwards of 256 station units and can have as many as 7170 stations by tying more than one exchange together.

Because they are digital, it is easy to set them up with quick dial single or double digit dialing of often used numbers, connect them with outside C/O or PBX lines, and record a customized outgoing message for callers and record their response, retrievable at any master. Because they are telephone system compatible, they are capable of supporting a DTMF generating, single line telephone instrument (telephone, autodialer, fax machine, etc.). Interconnections are through twisted pairs of copper wire or through fiber-optic cable.

Many of these systems are programmed with a computer. Through simple programming, only selected stations can be given access to special functions—i.e., paging, telephone calls, priority calls, external sources, etc. Compression circuits and/or automatic leveling circuits can be adjusted for individual units as well as their overall level.

43.1.3 Apartment Security/Intercom Systems

Apartment systems are usually point-to-point and are used for the visitor to contact the apartment owner for access into the building and then into the apartment. This system consists of a master panel outside the main access door to the building and slaves in each apartment, Fig. 43-4.

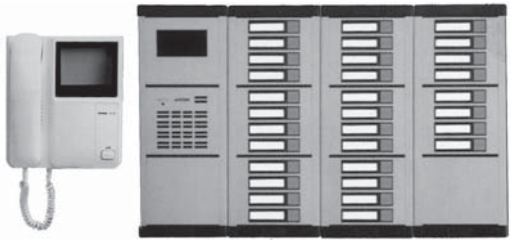


Figure 43-4. Audio/video entry security system. Courtesy Aiphone Communications.

The outside master panel must be waterproof and vandal proof so the outside unit panels and buttons must be made of strong material such as aluminum, stainless steel, or LEXAN®. The microphone/loudspeaker grill must be indestructible and shed water, preferably a continuous part of the panel and the loudspeaker must be waterproof. The panel must be fastened to the wall with vandal-free mounting hardware. The outside master panel consists of a microphone/loudspeaker and a means of contacting the person of choice. This can be accomplished by a series of push buttons that connect the visitor to the desired apartment or it may be a twelve button telephone panel that connects the caller to the suite through digital circuitry.

A power supply and amplifier is mounted inside where it is out of the elements. Only one amplifier and power supply is normally used for a system because only one conversation goes on at the same time. The door opener usually operates on 3–6 Vdc or 8–16 Vac.

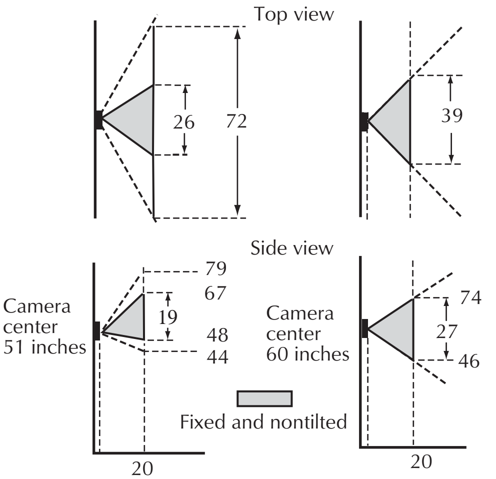
The apartment loudspeaker station consists of a microphone/loudspeaker unit, a talk button, a listen button, and a door release button.

The more sophisticated security systems include a video monitor for improved safety. The outside door panel includes a camera and the indoor units include a 4 inch flat screen monitor. A pan and tilt capability improves security as it can scan a large area. Because the camera is behind a glass, the visitor cannot see where it is aimed so the visitor cannot hide from it. With today’s technology, wide angle cameras are possible, eliminating the need for the more expensive and complicated pan and tilt. The cameras require very good low light level operation as most access is in the evening or at night. One lux sensitivity is normal. Most systems do not require coaxial cable between the cameras and the monitors.

Self-contained video door answering systems are rapidly becoming today’s replacement for the common doorbell. They are a simple-to-install answer to the growing need for entry security in both small businesses and homes. Some use the same two wires as a doorbell, often using existing wiring in older buildings, and simplifying installation in new construction.

An example is a system where one pair carries an FM modulated signal with both audio and video, and a second pair operates the door releases shown in Fig. 43-4. In this system the door unit incorporates a high resolution infrared CCD (charge-coupled device) camera with 250,000 pixels, providing a clear, sharp picture down to one lux of light.

By using a wide angle lens, an overall viewing area of 39 inches × 27 inches at 20 inches away can be attained, Fig. 43-5. For more coverage, a PanTilt door station can be used, Fig. 43-5. This allows for a coverage of 72 inches × 36 inches at 20 inches. Like any transmission system, cabling is important. The system was designed to operate with two wires such as typical bell wire. Coaxial cable or two separate wires or multicable will affect picture quality. If the cable has more than one pair, the other pairs should be terminated at each end with a 120 Ω resistor. This keeps a proper impedance of the line.



A. Pan and tilt. B. Wide angle lens.

Figure 43-5. Video coverage door stations. Courtesy Aiphone Communications.

Any door entry system camera should be mounted so the unit is not exposed to environmental extremes. Sudden temperature drops, for instance, can cause the camera to fog up and raindrops can distort the picture. Lighting should be on the front of the person. Strong backlight—i.e., sunlight or street lights—can cause a silhouette image on the monitor, and fluorescent lights can cause flickering. Taking these into account will give a good picture under most conditions.

It is important that the door release button has normally open contacts to activate an optional electric door strike. The normally open contact assures that the door will not open during a power failure.

43.1.4 Residential Intercom Systems

Residential intercom systems are used to talk between rooms, or between a master unit and all other rooms, as a door security system and for programming external sources such as AM/FM or CDs to any or all rooms. Because of the limitations of intercom systems, the music is usually not in stereo and not of the quality of a dedicated stereo system, but is quite adequate for background music. They also have the advantage that the same source is heard as a person walks between rooms. Normally residential intercoms are wall mounted, incorporate hands-free answering and include a privacy switch. The systems can be master/slave, master/master/slave, or all master.

43.1.5 Commercial Security Systems

A building security system assists security personnel in the protection of the lives and property of all tenants, employees, and visitors. People in parking areas, ramps, tunnels, stairwells, and elevators should have access to conveniently located, easily operated hands-free emergency call-in stations.

43.1.5.1 Zone Paging

High-rise buildings and multibuilding complexes have special needs for zoned public address announcements. Buildings with controlled access need audio voice confirmation and integration with CCTV cameras. Elevators require intercommunication to security and the lobby and communication to the elevator machine room for maintenance, Fig. 43-6.

Variations of this type of intercom can be used for campus security where there are multiple buildings, parking lots, dorms, and walkways.

43.1.5.2 Security Audio Monitoring

Unfortunately, all too often crisis situations do develop in the shadows of darkened areas or dimly lit passageways. These are the areas where criminals tend to stalk. These are also the places where a building's highly volatile power transformers, generators, and steam lines are neatly tucked away. Security and maintenance personnel can't be everywhere. Even with the assistance of video surveillance there are limits to the number of cameras used and manpower to monitor them. Even in the best of situations video cameras can't see around corners, through closed doors, or behind parked cars or trucks.

To eliminate problems with video only monitoring, security system manufacturers have developed listening devices and loud-noise triggering alarms. Although these products seem to be moving security systems in the right direction, they still have substantial limitations.

Listening devices are a great idea because they allow security personnel to interpret and discriminate the sounds they hear. Unfortunately, like video cameras, they must continually be monitored to be truly effective. There is also the question of open microphones being construed as invading individual privacy.

Loud-noise triggering alarms have the problem of discerning sounds, for instance, screams from laughter, or the loud noise of a car engine starting up. Without the ability to discern specific sounds and discard normal background sounds, false alarms would render the system useless.

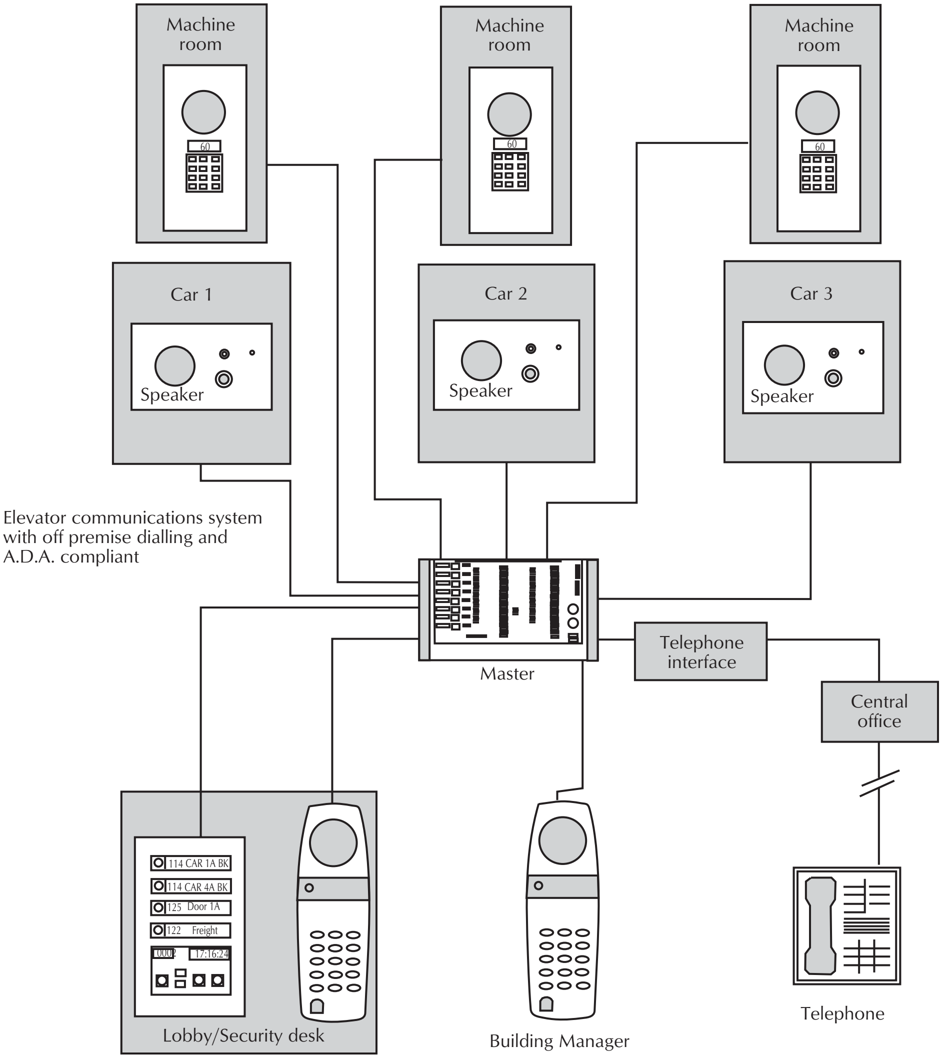

Machine room
60
Machine room
60
Machine room
60
Car 1
Speaker
Car 2
Speaker
Car 3
Speaker
Elevator communications system with off premise dialling and A.D.A. compliant
Master
Telephone interface
Central office
114 CAR 1A BK
114 CAR 4A BK
125 Door 1A
122 Freight
0002
17:16:24
Lobby/Security desk
Building Manager
Telephone

Figure 43-6. Basic building communication system. Courtesy Ring Communications, Inc.

Today there are *smart* security devices that incorporate the best features of both a listening device and a noise trigger without many of the drawbacks.

One effective method is a sound sensing device that continually adjusts itself to background noise levels, while detecting and discerning the unique sound characteristics of specific sound patterns usually associated with crisis and emergency situations. When the system picks up one of the crisis sound patterns it fires and alerts security personnel of a probable emergency situation in a particular area. The system can also automatically turn on video cameras, sound alarms, or activate two-way intercoms, so that security personnel can instantaneously communicate with individuals involved at the scene.

Because this type of system continually adjusted itself to the continuous changes in levels and frequency response of background sounds, and has the ability to insert a time gate to further assist in doing away with false alarms, it is able to isolate the sound of a scream or the smashing of a pane of glass, Fig. 43-7.

This type of monitoring is useful in security applications including correctional facilities and mass transit subway authorities, parking lots, schools, large warehouses, and empty office building hallways after regular business hours.

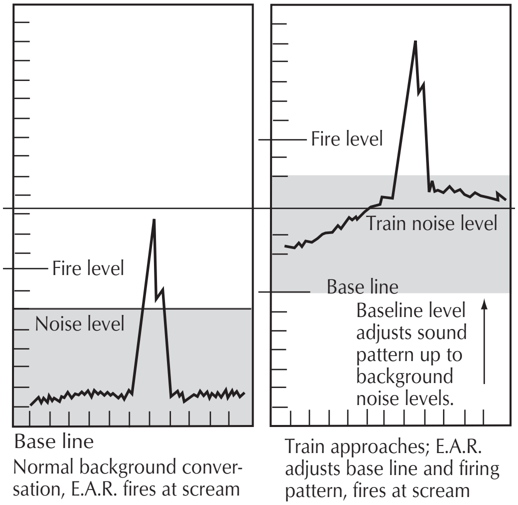


Figure 43-7. Simulated firing of E.A.R. in an active subway station. Courtesy Ring Communications, Inc.

43.1.5.3 Emergency Crisis Communications

Methods of communications change, methods of transmission change, and methods of installation change. Manufacturers continue to upgrade their equipment to handle these changes.

At one time the basic accepted method of signal transmission was via copper conductors only, now there are multiple methods to accomplish this.

Voice communications is often overlooked in some areas, but an alert that is only a bell or horn or chime sounding only alerts people that there is a problem, it does not tell them what the problem is, where it is, or what response they should take to protect themselves. By adding voice communications to panic alarms, fire alarms, or crash alarms, the general public can be advised of what the nature of the alarm is, where it is, and what to do to avoid it.

Often huge geographical areas must be covered. To do this with copper alone is almost impossible. The Ring Master Intercommunications Crisis Alert System has the ability to tie equipment together to operate as one complete system, providing for line supervision over multiple methods of transmission, simultaneously solving any such problem.

Fig. 43-8 shows such a system that is in use today to cover the many and varied needs of an airport site.

The system provides communications for all aspects of airport communications requirements over one integrated supervised system. The system includes:

- Air traffic control.
- Security.
- Baggage handling.
- Staff location.
- Access to general over-head paging.
- De-icing.
- Flight operations and planning.
- ADA elevator communications.
- Door access and control.
- Emergency call boxes for the parking facilities.

When distances between buildings and operational sites are long, copper is not a viable solution. Fiber is better utilized to tie geographically distant areas together. In other areas the newest method of VOIP, over radio links, can be utilized to provide communications, for instance, to the people mover train system.

With this type of system, any station in the system, no matter where its geographical location on the airport site, can be programmed remotely, to call or be called by any other station in the system or to access any of the multiple functions and features of the communications system at any time.

When these connection capabilities are added to the commercial security intercom system there are virtually no restriction to fulfill the requirements for safe, secure communications.

43.1.6 Commercial/Industrial Systems

Commercial/industrial intercom systems are used in airports, car dealers, dormitories, factories, hospitals, nursing homes, department stores, and schools. Most of these systems are master/slave systems or subsystems.

The health care and hospital facilities systems must be designed for emergency hands-free operation from operating rooms, trauma rooms, delivery rooms, etc. The systems must also be capable of sterilization after each use. This can be accomplished with a Mylar covered face plate that is easy to sterilize because of lack of crevices, etc. The system can include a paging system for doctor, nurse, or security call and for calling patient's in waiting rooms.

Operations of general commercial/industrial systems are very much like the door access or residential system but include many of the following special features:

Absence Registration. A unit can be programmed to display the station is unattended.

All Call. All call allows a master station to connect to every other station for an all-call message.

Call Back. Call back is initiated by the caller when the receiving line is busy. The caller can put the system on automatic call back which notifies him as soon as the

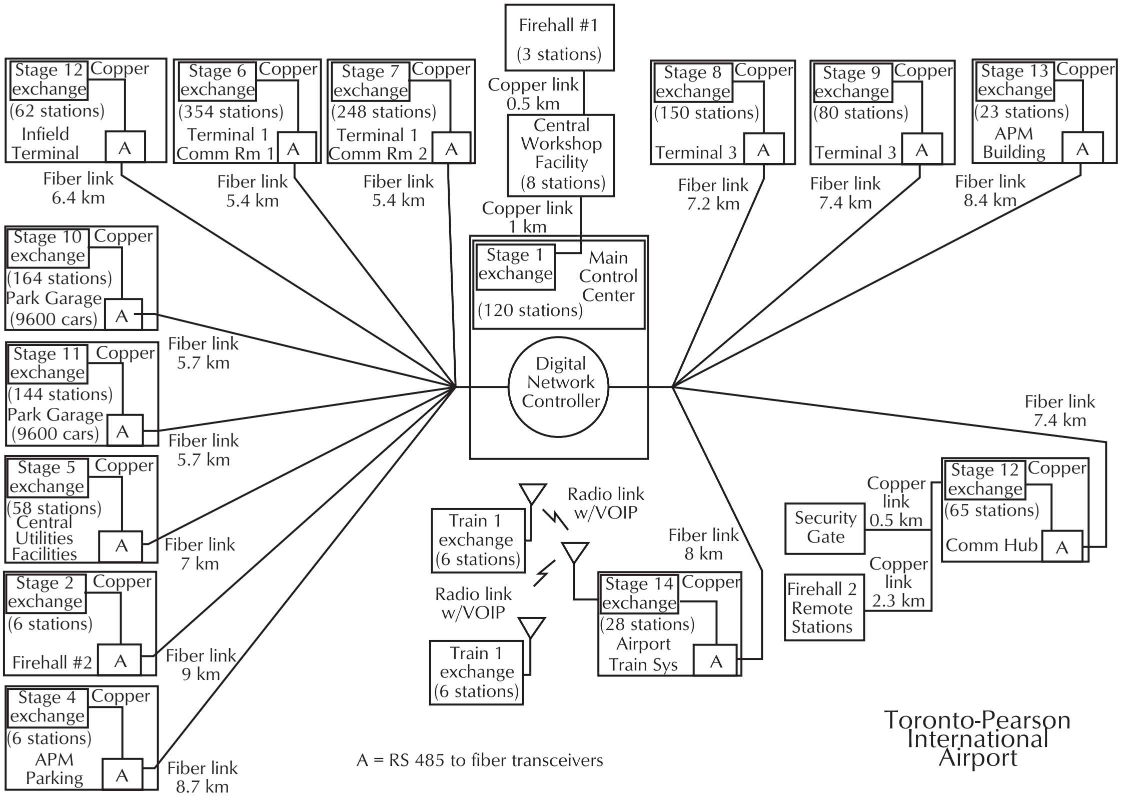

Firehall #1
(3 stations)
Copper link
0.5 km
Central
Workshop
Facility
(8 stations)
Copper link
1 km
Stage 1
exchange
(120 stations)
Main
Control
Center
Digital
Network
Controller
Stage 12
exchange
Copper
(62 stations)
Infield
Terminal
A
Fiber link
6.4 km
Stage 6
exchange
Copper
(354 stations)
Terminal 1
Comm Rm 1
A
Fiber link
5.4 km
Stage 7
exchange
Copper
(248 stations)
Terminal 1
Comm Rm 2
A
Fiber link
5.4 km
Stage 8
exchange
Copper
(150 stations)
Terminal 3
A
Fiber link
7.2 km
Stage 9
exchange
Copper
(80 stations)
Terminal 3
A
Fiber link
7.4 km
Stage 13
exchange
Copper
(23 stations)
APM
Building
A
Fiber link
8.4 km
Stage 10
exchange
Copper
(164 stations)
Park Garage
(9600 cars)
A
Fiber link
5.7 km
Stage 11
exchange
Copper
(144 stations)
Park Garage
(9600 cars)
A
Fiber link
5.7 km
Stage 5
exchange
Copper
(58 stations)
Central
Utilities
Facilities
A
Fiber link
7 km
Stage 2
exchange
Copper
(6 stations)
Firehall #2
A
Fiber link
9 km
Stage 4
exchange
Copper
(6 stations)
APM
Parking
A
Fiber link
8.7 km
Train 1
exchange
(6 stations)
Radio link
w/VOIP
Radio link
w/VOIP
Train 1
exchange
(6 stations)
Fiber link
8 km
Stage 14
exchange
Copper
(28 stations)
Airport
Train Sys
A
Fiber link
7.4 km
Security
Gate
Copper
link
0.5 km
Firehall 2
Remote
Stations
Copper
link
2.3 km
Stage 12
exchange
Copper
(65 stations)
Comm Hub
A
A = RS 485 to fiber transceivers
Toronto-Pearson
International
Airport

Figure 43-8. Ring Master Intercommunications Crisis Alert System at Toronto-Pearson International Airport. Courtesy of Ring Communications, Inc.

intended receiver is free so he can reinstate the call or the system can automatically recall the receiver.

Call Forwarding. Call forwarding allows all calls directed to a station to be forwarded to another station.

Call Reply. When the called person is absent, the caller may dial a call-back signal which indicates that he has been called.

Call Transfer. Particularly useful for transferring a call to or from a secretary.

Camp On. Camp-on is a feature which allows the caller to camp-on to the called person when that person is in communication with someone else. Once the called person hangs up, the caller is automatically connected to him. Normally, camp-on only holds on for 10–20 s before dropping out, or may remain in service until the next call is initiated.

Central Answering Service. Units can be set up to go through a central answering person with automatic queuing of incoming calls.

Conference. Conference calls can be held with more than one party. Normally three to four parties is the maximum limit, however some systems may be expanded to a 30 party conference.

Display Features. Today's intercoms usually use an LCD type display system which gives the called or calling station number, in addition may display a series of preprogrammed messages, whether or not the called station is in privacy mode and time/date.

Group Callback. Many systems are setup so that various groups can be connected with one button so a person can call an entire group at one time.

Hands Free. In this condition, conversations are hands free from anywhere in the room as long as it is a duplex system. This can be changed to confidential at any time by a person picking up the handset and operating as a telephone system.

Hurry Up. If the conversation channel is busy, and the caller needs to make an emergency or an important call, she can transmit a special signal on the conversation channel which tells them there is an important message to be received.

Last Number Redial. Last number redial, as on telephone systems, allows the caller to redial his last called number again and again and again.

Microphone Cutout. Microphone cutout is used to disconnect the microphone during a conversation so that the receiver cannot hear communications between the talker and another person in the same room.

Paging. Master stations can page to other master stations, substations, and to remote loudspeakers through an external amplifier.

PBX Station Intercommunications. Many intercoms can be connected to the PBX or an outside C/O line for communications to the outside world.

Pocket Page Access. Any master station can place calls to the desired pocket pager receiver, up to 10,000 units.

Priority. If a priority station dials a number while the conversation channel is busy, she has the option to override the busy link and be connected directly through to the called party. This is normally used for emergency systems.

Privacy. Privacy is initiated by a person so that a caller, when calling that person, hears a signal which tells him the receiving person does not want to be disturbed.

Program Distribution Channels. Music or special messages can be setup to be used as background music.

Scan Monitoring. Scan monitoring allows a control station to arbitrarily scan a group of slaves or substations for auditory monitoring. Scanning can be performed either manually or automatically.

Time/Date. Some intercoms have the capability of displaying the time and the date.

Nurse Call Systems. Nurse call systems must be fast, reliable, and most of all easy to use. Patients must not have to figure out how to use a system to call for help. The system must not interfere with the complicated patient monitoring systems and must not produce ground loops between it and other equipment.

The simplest and probably best method for a patient to call a nurse is with a pull chain. This chain, actually cord, electrically isolates the system from the patient and only requires a gentle pull by the patient to operate. The cord can be draped on the rail of the bed for easy access by the patient. A monitoring system is located at the nurse's station with lights corresponding to room numbers or it could be a CRT with various messages. An indicator light outside the patient's room is also energized, allowing a nurse in the hallway to answer a call without going back to the nurse's station. The call remains energized until the nurse turns it off from the patient's station. Because the nurse call systems only tell the nurse which room is calling, dual-bed rooms have one call station with two call cords.

Nurse call systems can also include master/slave two-way communications, eliminating a nurse's extra steps. It also allows a nurse station to monitor a room. Some areas require an emergency button. This could be a pull chain, a large red mushroom button, or in the case of a psychiatric station, it could be key operated.

Equipment Locator. An equipment tracer system such as the Rauland Borg Tracer system saves hospitals time, money, and frustration by locating their key people and equipment quickly and quietly.

Using a Responder III Plus Nurse Console, staff can instantly locate and communicate with key staff or locate equipment. Equipment and staff locations can be displayed in real time on networked PCs throughout the facility.

Tracer works in the following manner. Three elements comprise the system: tracer tags, in-ceiling network, and software elements. The Windows-based software gives a graphic, easy-to-understand location depiction or a synthetic voice can give location over any telephone.

Special light-weight transmitter tags can be attached to a facility's important equipment (IV pumps, wheelchairs, carts) and worn by key personnel. These tags continuously relay their position to room and hallway sensors, Fig. 43-9.

The tag is microprocessor based and emits an 880 nm infrared hemisphere of one of 30,000 unique ID digitally encoded packets for a broadcast range of about 20 ft.

Ceiling or wall-mounted sensors receive the packets and convert them tzo electrical signals. The packets are relayed to the collector box, which has star inputs for twenty four sensors.

The collector assembles the sensor packets into larger packets and prepares them to be sent to the concentrator.

Each senso, or combination of sensors, is defined as a zone. Offices, hallways, or areas can be declared as zones.

Data from the network (tag, location) is combined with the time of day and written to binary history files, called *Logger*. The database and history files can be

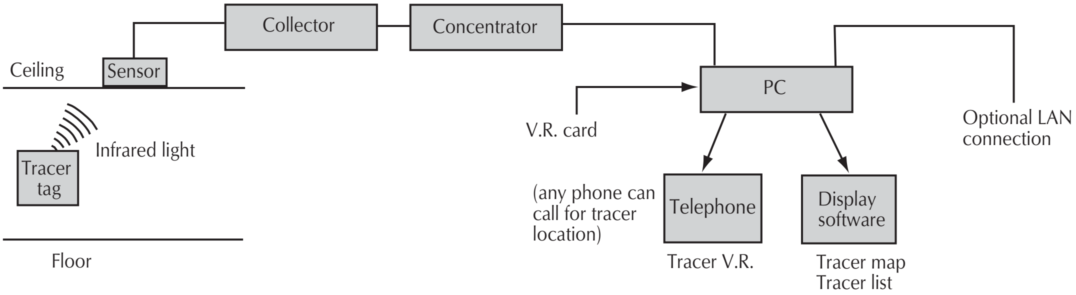


Figure 43-9. Hardware elements of a tracer system.

retrieved from the foundation of tracer software. The multitasking Microsoft Windows operating system is used, which creates an open system.

Three presentation methods are available. TracerLIST handles multiple floors and hundreds of tags. Features include a living directory presentation, fast alphabetic search, usual location and extension, current location, occupants currently in the room, and system condition heartbeat.

Tracer VR (voice response) uses a voice response card and allows users to dial into the system to audibly learn location. Tracer VR extends the tracer capability to before and after-hours callers, and personnel without workstations.

TracerMap is a living floor plan that gives a graphical location of Tags within an area.

43.1.7 Wireless Intercoms

When areas are constantly being reorganized, a wireless intercom may be useful. A wireless intercom may use RF for transmitting the signal or it may be transmitted through the 120 Vac lines. When the transmitting system is RF, two frequencies are required per unit, one for talking and one for listening unless the system operates like walkie talkies where simultaneous two-way communication is not possible. If each unit must have its own private communication, then each unit would require a different transmit and receive frequency or use a system of subaudio modulation to key the appropriate unit. The subaudio modulation system can have only one conversation at a time. The biggest drawbacks of wireless intercoms is their ability to pick up noise and stray signals and fewer bells and whistles are available. However, if the area is confined and relatively free from electrical noise, wireless intercoms can greatly reduce installation costs.

When the ac power lines are used to connect the audio between units, the frequency response must be limited to reduce noise and hum. If the two units are on different legs of the two phase 220/110 Vac transformer, communications may not be possible without some means to transfer the audio signal between the legs.

43.2 Broadcast Intercoms

The need for rapid, reliable, and flexible communications is required in broadcast intercoms. Fortunately, intercom equipment is capable of meeting almost any need that might arise.

Telephones are usually used for less than 20 minutes per call. Intercoms are often used for many hours at a time. The people involved in teleproduction work often can't take a break or remove their headsets. If the system has limited frequency response, then the system's filter effects create distortion. This unnatural sound can cause fatigue, which can be eliminated with a full-frequency intercom.

Broadcast intercoms can be point-to-point, party line, or matrix. The audio line can be balanced or unbalanced. Balanced line operation provides maximum protection from electromagnetic interference generated from sources such as fluorescent lights, patch panels, or light dimmers. The 24 Vdc phantom power is fed down the audio line with the minus voltage on the shield and the positive voltage bridged between the two shielded lines. The bridging is accomplished by connecting two resistors in series and across the balanced line and connecting the positive voltage to the junction between the resistors. It is important that the resistors be at least 1% tolerance or better. A balanced system can run as much as 5000 ft with standard two-conductor shielded microphone cable.

An unbalanced line, sometimes called a three-wire system, is easier to switch and operate special circuits as the audio and signal are on different lines, only using the ground (shield) as a common line. A second method provides two channels of unbalanced audio, one channel between each conductor and shield, and combining the dc operating power with one of the audio channels.

Many intercoms are engineered to receive phantom power from the 24 Vdc system power supply. Stationary or permanent user stations usually operate in a dry mode as they are supplied with power from a local power line. The term *dry* refers to an intercom channel that has audio but not the usual 24 Vdc phantom power on the channel. Dry operation has several advantages over wet operation as it is generally quieter, reduces the need and cost of large system power supplies, and takes up less master rack space. System configurations can include a mix of wet and dry channels, depending on the station equipment assigned to the particular channel. Generally, most wired belt-pack and loudspeaker stations require a wet channel and thus need a system power supply.

43.2.1 Broadcast Point-to-Point Systems

The point-to-point system consists of a centralized rack (or racks) of amplifiers and signal routing circuitry controlled from remote stations. The audio signal paths are analog and simplex or digital.

The system allows a station to route its voice or other signal to one or more other stations. The originating talker determines who hears the communication and the listener normally has no control over who is received at the individual stations. Normally, point-to-point systems require direct or home run cabling from each station to the central control.

Each station in a point-to-point analog system requires a minimum of one audio *transmit* pair, one *receive* pair, many control or station selection conductors, and a power pair or local power supply. It is essentially a switch-selected, multiple station, one-way paging system. A point-to-point digital system has the same functions plus many more and operates on a single pair of unshielded wires.

43.2.2 Broadcast Party Line System

In the party line system, each station is equipped with all of the required electronics for receiving and transmitting audio and for signal routing. Party line systems require minimal centralized rack equipment, which usually consists of the system power supplies and passive assignment switching in multichannel systems.

Party line systems allow groups of stations to communicate in real-time, full duplex fashion. In fact as it is a party line, all units on the line hear and are part of all conversations. Multichannel party line systems allow users access to several different channels, allowing them to determine which line they talk and listen to. Normally there are no private communications such as point-to-point systems provide.

Two-channel, dual-listen, with monaural output intercoms with programmable switching let the user listen to both channels simultaneously and select which channel to talk on. They include an individual volume control for each channel, microphone on/off control, a call signal button and indicators, and sidetone. These stations are ideal for ENG and EFP mobile production vans, production studio consoles, and TV facilities.

Straight two-channel units allows simultaneous listening and talking on two intercom channels. The headphone output operates in a split-feed stereo mode, feeding each channel into a separate ear of a double-muff headset with an individual volume control for each channel. The operator can talk or listen on either channel, combine them, or access them separately without tying both channels together. Often a stage announce output can be supplied with relay control for external paging. Microphone or line-level program inputs may be assigned to either or both channels. The systems capabilities can include a selectable *program interrupt* function, remote m*ic-kill* function, dual-action, electronic momentary/latching *talk* buttons, and a n*o-fail* power supply with automatic reset and short-circuit protection.

A typical two-channel party-line system by Clear-Com consists of a main station, which includes a power supply. Each channel is full duplex allowing everyone on the channel to choose to both listen and talk to everyone else. Although the min station provides two channels, single- and two-channel beltpacks can be mixed within the same system. This reduces cost and allows one to control exactly who is assigned to what channel. All cabling is standard low capacitance shielded microphone audio wiring, Fig. 43-10. A party line belt pack is shown in Fig. 43-11.

Some multichannel systems have extensive programming capabilitie, which allows individual stations to be customized by storing the *button setups* in nonvolatile memory. Many programmed setups can be stored, thus allowing quick and easy switching between setups for rehearsal and performance or shows and events. Individual button assignments can be stored in *presets* for instantaneous recall. When programming this equipment, messages prompt the operator through the programming sequence, simplifying station setup.

A four-channel, two-wire party line system is similar to the two-channel party line except with two additional channels. In Fig. 43-12, Channel A is for the floor crew, Channel B is lighting, Channel C serves audio with Channel D calling the talent dressing rooms and

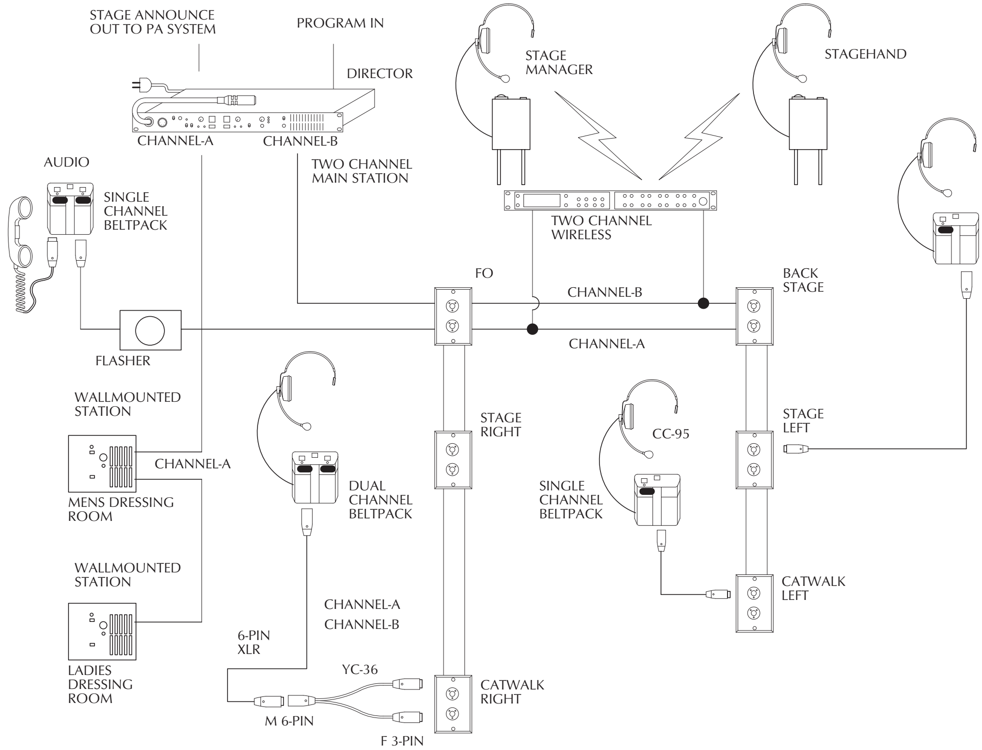

STAGE ANNOUNCE OUT TO PA SYSTEM
PROGRAM IN
DIRECTOR
STAGE MANAGER
STAGEHAND
CHANNEL-A
CHANNEL-B
TWO CHANNEL MAIN STATION
AUDIO
SINGLE CHANNEL BELTPACK
TWO CHANNEL WIRELESS
FO
BACK STAGE
CHANNEL-B
CHANNEL-A
FLASHER
WALLMOUNTED STATION
STAGE RIGHT
STAGE LEFT
CC-95
CHANNEL-A
MENS DRESSING ROOM
DUAL CHANNEL BELTPACK
SINGLE CHANNEL BELTPACK
WALLMOUNTED STATION
CATWALK LEFT
CHANNEL-A
CHANNEL-B
6-PIN XLR
LADIES DRESSING ROOM
YC-36
CATWALK RIGHT
M 6-PIN
F 3-PIN

Figure 43-10. Two-channel party line system. Courtesy Clear-Com Communication Systems.

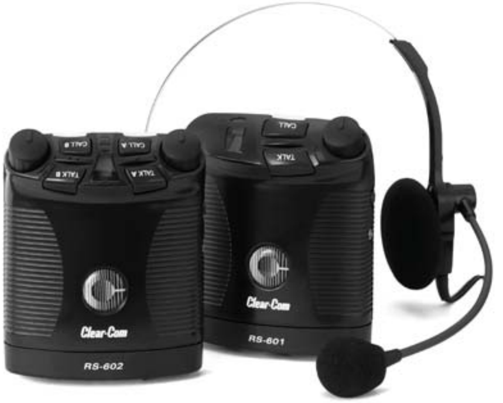

CALL B
CALL A
TALK B
TALK A
Clear-Com
RS-602
CALL
TALK
Clear-Com
RS-601

Figure 43-11. A party line belt pack. Courtesy Clear-Com Communication Systems.

box office. The producer and director have access to all four channels. The lighting director is limited to the floor and light crew. The audio engineer can access the floor crew and channel D audio. Note the majority of the belt packs is single channel. There is no reason for the lighting crew to talk to the dressing rooms, which keeps conversations specific to their assigned channels. All cabling is standard low capacitance shielded microphone audio wiring.

43.2.3 Broadcast Matrix Systems

Matrix systems are a powerful and cost-effective communications tool. Expansion is easy and installation is inexpensive as only a single pair of wires is required for audio, signaling, and external inputs. System size ranges from 2 × 2 to over 784 × 784.

There are two basic methods of interconnection between matrix intercom stations and the central frame: a*nalog audio* and *fully digital*.

The two-wire party line matrix system is a very flexible multi-channel party line where numerous sources, or drops, are connected to an 8ft × 24ft assignment panel. The drops can be mixed. The example of Fig. 43-13 reflects twenty four drops across eight PL channels. The configuration permits the user to program any one of the twenty four locations to a specific PL channel. For example, one day you might need a camera position at stage left and tomorrow the lighting crew might use that location. You may also route wireless intercom, two-way radios, and fiber remote links to any

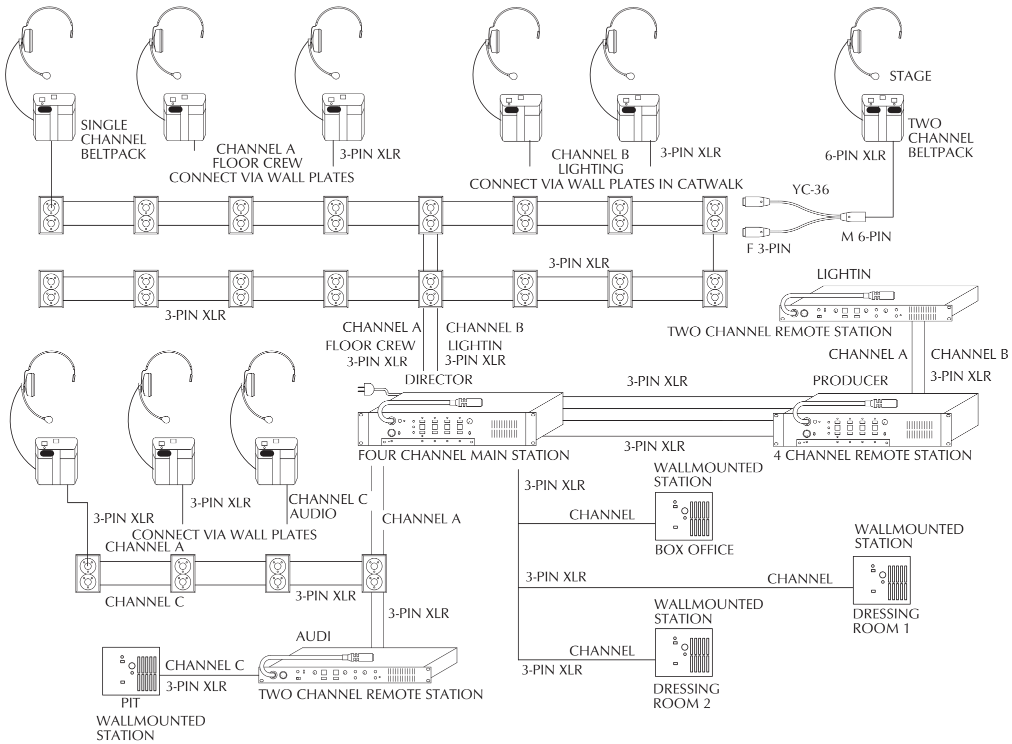

STAGE
SINGLE
CHANNEL
BELTPACK
CHANNEL A
FLOOR CREW
CONNECT VIA WALL PLATES
3-PIN XLR
CHANNEL B
LIGHTING
3-PIN XLR
CONNECT VIA WALL PLATES IN CATWALK
6-PIN XLR
TWO
CHANNEL
BELTPACK
YC-36
M 6-PIN
F 3-PIN
3-PIN XLR
3-PIN XLR
LIGHTIN
TWO CHANNEL REMOTE STATION
CHANNEL A
FLOOR CREW
3-PIN XLR
CHANNEL B
LIGHTIN
3-PIN XLR
CHANNEL A
CHANNEL B
3-PIN XLR
DIRECTOR
3-PIN XLR
PRODUCER
FOUR CHANNEL MAIN STATION
3-PIN XLR
4 CHANNEL REMOTE STATION
WALLMOUNTED
STATION
3-PIN XLR
CHANNEL
BOX OFFICE
3-PIN XLR
3-PIN XLR
CHANNEL C
AUDIO
CHANNEL A
CONNECT VIA WALL PLATES
CHANNEL A
WALLMOUNTED
STATION
3-PIN XLR
CHANNEL
DRESSING
ROOM 1
CHANNEL C
3-PIN XLR
3-PIN XLR
WALLMOUNTED
STATION
AUDI
CHANNEL
3-PIN XLR
CHANNEL C
3-PIN XLR
TWO CHANNEL REMOTE STATION
DRESSING
ROOM 2
PIT
WALLMOUNTED
STATION

Figure 43-12. Four-channel, two-wire party line system. Courtesy Clear-Com Communication Systems.

of the eight traditional party line channels. Fourteen pre-programmed configurations may be stored and recalled on the fly. All cabling is standard low capacitance shielded microphone audio wiring.

The standard analog wiring format provides wide audio bandwidth and easy connection over standard four-wire audio and data circuits. It allows station-to-matrix communication through repeaters and over satellite links, ISDN circuits, and fiber optic systems.

The fully digital format uses only a single pair of wires that carry data plus digitized audio, simplifying installation. Additionally, digital audio cannot be tapped by unauthorized persons, it offers total noise immunity from external sources. The wiring is easy, it permits standard punch-block connection techniques using existing unshielded multipair telco or CAT 5 wiring.

The digital matrix system differs from Party Line in several areas. In addition to party line conversations, one may also conduct private point-to-point conversations between the panels. Through intuitive programming, members of the groups may be changed at will. This system easily integrates telephones, two-way radios, line-level audio in/out, voice over IP, GPIs, and relays, which allows incoming and outgoing telephone calls to be routed to particular panels or groups. Matrix systems are well suited for dynamic broadcast environments where IBF audio feeds must change quickly. A wireless intercom is easily integrated and can be treated as another panel depending on the wireless version in use. The connections between the mainframe and control panels are standard nonshielded Cat 5 or 6 cables. A digital matrix system by Clear-Com is shown in Fig. 43-14.

Matrix systems use crosspoint and CPU circuitry. The mainframe serves as the central interconnection point for the control stations, interface modules, power supplies, the configuration computer, and external audio and control equipment. All signals, digital and analog, are processed in the mainframe and routed according to the current software configuration program.

The CPU operates the frames and control all of the system data communications. Crosspoint electronics contain microprocessors that communicate with the CPU and with the stations. The crosspoint circuitry supports individual ports that can connect to stations, interfaces, or to analog four-wire circuits and devices.

Interface circuitry couples the master system to external four-wire circuits providing the proper isola-

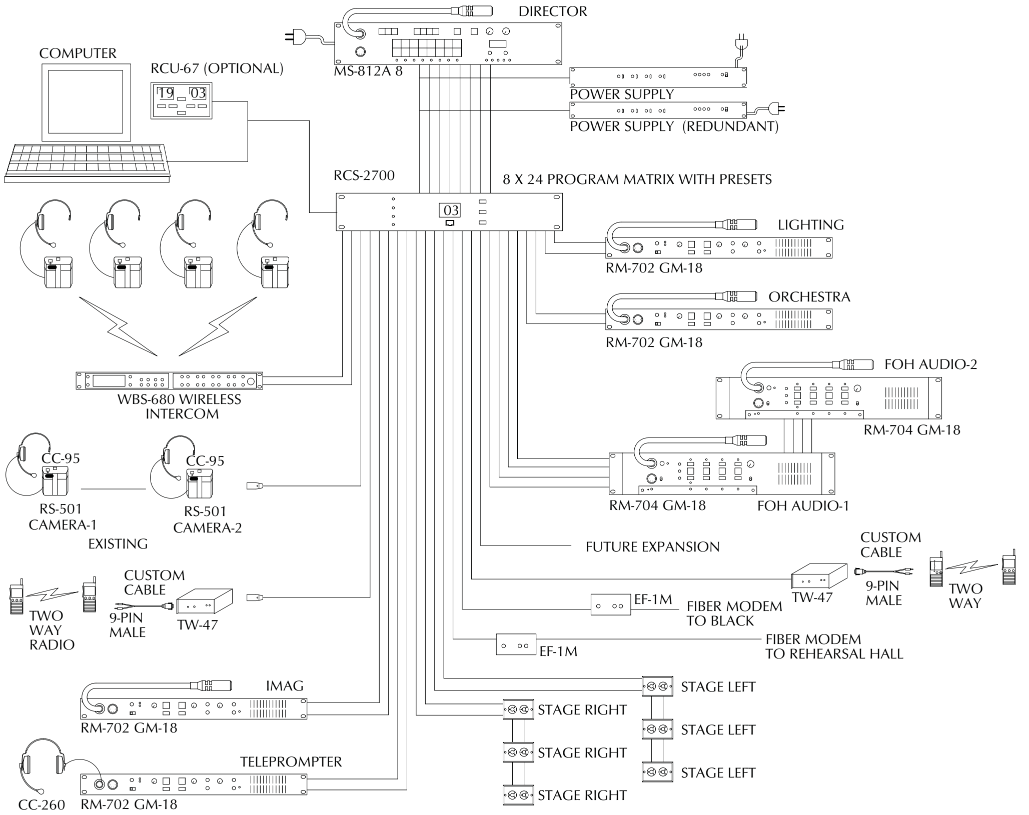

DIRECTOR
COMPUTER
RCU-67 (OPTIONAL)
19 03
MS-812A 8
POWER SUPPLY
POWER SUPPLY (REDUNDANT)
RCS-2700
8 X 24 PROGRAM MATRIX WITH PRESETS
03
LIGHTING
RM-702 GM-18
ORCHESTRA
RM-702 GM-18
FOH AUDIO-2
RM-704 GM-18
WBS-680 WIRELESS INTERCOM
CC-95
CC-95
RS-501 CAMERA-1
RS-501 CAMERA-2
EXISTING
RM-704 GM-18
FOH AUDIO-1
FUTURE EXPANSION
CUSTOM CABLE
TW-47
9-PIN MALE
TWO WAY
CUSTOM CABLE
TWO WAY RADIO
9-PIN MALE
TW-47
EF-1M
FIBER MODEM TO BLACK
EF-1M
FIBER MODEM TO REHEARSAL HALL
IMAG
RM-702 GM-18
STAGE LEFT
STAGE RIGHT
STAGE LEFT
STAGE RIGHT
STAGE LEFT
STAGE RIGHT
TELEPROMPTER
CC-260
RM-702 GM-18

Figure 43-13. 8 × 24 program two-wire party line matrix with presets. Courtesy Clear-Com Communication Systems.

tion, impedance matching, and level sets between systems. Additionally, it supports external relay activation and call sense circuitry. Typical circuits include 4-wire telephone lines, camera intercoms, two-way radios and microwave links, four-wire intercoms, fiber optic lines, and satellite links. Other types of four-wire circuits include IFB systems, ISO systems, and program audio in/out.

Many matrix systems have the capability of linking to other matrix systems. This allows stations in one system to communicate with stations in other systems and anything that can be selected or controlled in one system can be selected and controlled from any other linked system. This allows independent systems in remote locations, even in different cities, to operate a single system.

The station-to-matrix wiring in matrix systems can operate in either a three-pair mode or a four-pair mode. Using the four-pair scheme, remote station operation is possible from any location that can provide a standard four-wire audio circuit plus a four-wire RS-422 data circuit back to the central matrix. This can include transmission links such as satellite and fiber optic circuits, T-1 channel banks, ISDN, and switched 56 terminal adapters.

Port input signals are usually routed through software controllable *digital potentiometers*. This enables remote control of input levels from a PC, or directly from the intercom control stations, allowing instant on the fly adjustment of audio signals that are too loud or too soft.

Because the audio is digital, noise pickup is nonexistent, frequency response is 100 Hz–15 kHz ±1 dB, distortion is less than 0.5% ,and the SNR is greater than 60 dB.

43.2.4 Program Interrupt (IFB)

IFB is a television production trade acronym that stands for *interrupted feed-back*, *interrupted fold-back*, *interrupted return-feed,* and in some cases *prompt-mute*. IFB plays an important role in the behind-the-scenes activities that make a production, as it permits the director or

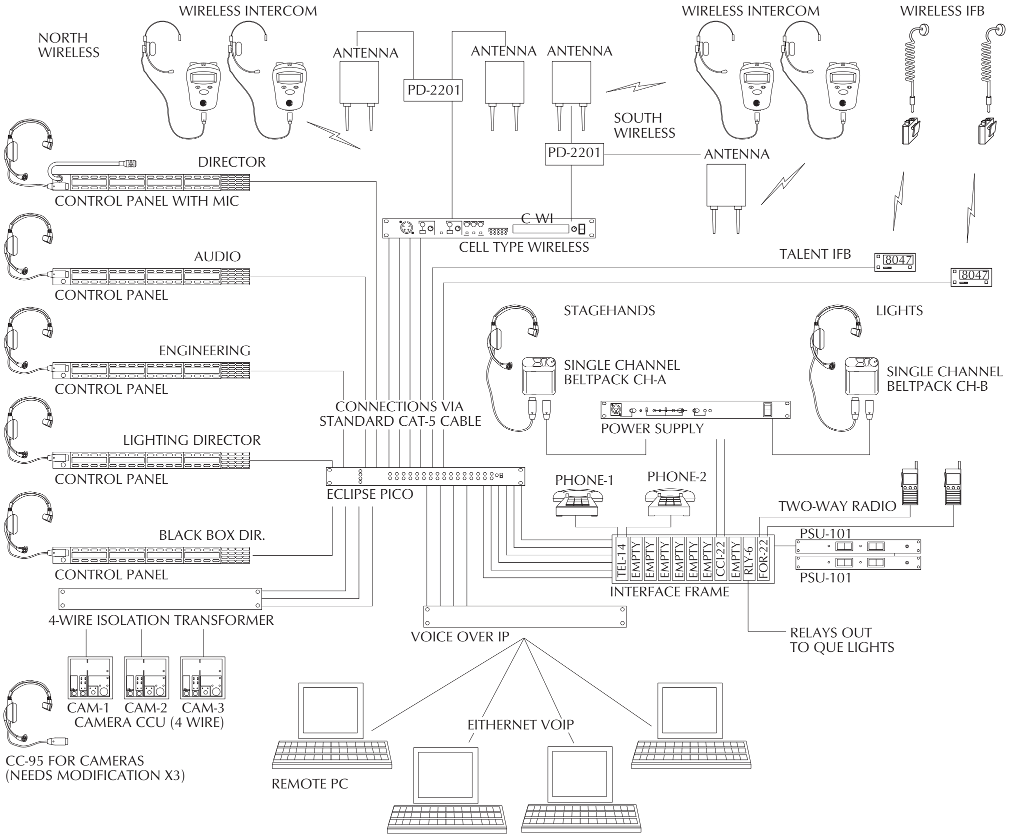

WIRELESS INTERCOM
NORTH WIRELESS
ANTENNA
PD-2201
ANTENNA
ANTENNA
WIRELESS INTERCOM
WIRELESS IFB
SOUTH WIRELESS
PD-2201
ANTENNA
DIRECTOR
CONTROL PANEL WITH MIC
C WI
CELL TYPE WIRELESS
AUDIO
CONTROL PANEL
TALENT IFB
8047
8047
STAGEHANDS
LIGHTS
ENGINEERING
CONTROL PANEL
SINGLE CHANNEL BELTPACK CH-A
SINGLE CHANNEL BELTPACK CH-B
CONNECTIONS VIA STANDARD CAT-5 CABLE
POWER SUPPLY
LIGHTING DIRECTOR
CONTROL PANEL
ECLIPSE PICO
PHONE-1
PHONE-2
TWO-WAY RADIO
BLACK BOX DIR.
CONTROL PANEL
PSU-101
TEL-14
EMPTY
EMPTY
EMPTY
EMPTY
EMPTY
CCI-22
EMPTY
RLY-6
FOR-22
PSU-101
INTERFACE FRAME
4-WIRE ISOLATION TRANSFORMER
VOICE OVER IP
RELAYS OUT TO QUE LIGHTS
CAM-1
CAM-2
CAM-3
CAMERA CCU (4 WIRE)
EITHERNET VOIP
CC-95 FOR CAMERAS (NEEDS MODIFICATION X3)
REMOTE PC

Figure 43-14. Digital matrix system. Courtesy Clear-Com Communication Systems.

producer to talk to the talent during a voice-over commentary either on or off camera in a live or taped production. Sports broadcasts typically use many channels of IFB to communicate with announcers in various locations on the field and in booths. When recording live performances on stage or in a studio, musicians can hear direct queues from the conductor or producer as well as the individual music mix designated for them. For television broadcasting, IFB queues on-camera announcers and is used between on-scene reporters and the in-studio anchor and studio director. IFB is controlled by the talker or person in control while the listener has no control over IFB.

43.2.5 Wireless Broadcast Intercom System

There are two types of wireless intercom systems. The first type provides a one-way, listen-only feature for the remote stations. In this system, the intercom is wired to the master transmitter. All communications on the intercom line are relayed to all of the wireless receivers. A one-way intercom is typically used for people who need to know what is going on, but don't need to talk back.

The second type of wireless intercom is a two-way system. In this configuration, the base unit and the field units can both talk and listen to each other in a full duplex mode. This requires two frequencies, a talk frequency and a listen frequency.

Wireless systems can stand alone, but when connected to a wired intercom system, the wireless link is virtually transparent to the user. The FCC has approved the 150–216 MHz band for broadcast use with over 1700 possible frequencies available. This band is relatively free from external radio and electrical interference. Transmitter output power is limited to 50 mW. Operating distances between units vary with the environment. If it is in open areas, it is possible to have good reception up to 2000 ft. However if the area includes walls, obstacles, and other radio transmitters, transmitting distances are more likely to be 150–300 ft. Battery life on belt packs should be over 20 h as these units are often used for extended periods. Long life is obtainable because the transmitters are only on when talking.

FM transmission has a characteristic that is called capture and is normally rated as capture ratio. It occurs when two or more transmitters are on the same frequency. The stronger of the two signals at the receiver captures the receiver, blanking out the other transmitters. If the transmitters are moving with respect to the receiver, the stronger transmitter would capture the receiver, so the communications could be bouncing between the various transmitters. For this reason all transmitters must be off when not in use and the new talker must monitor the channel before transmitting. If multiple transmitters and receivers are required, multiple frequencies are required.

43.2.6 Belt Packs

Belt packs are used in remote areas and when the person has to move around and/or the communications must not be heard by others such as during a theatrical performance, Fig. 43-15, or for football communications between spotters and coaches, Fig. 43-16. They may be wired or wireless and incorporate either single or dual headsets. Belt packs often utilize noiseless, digital electronic switching on audio circuits. A push-pull amplifier supplies high levels of audio in the headset and a microphone limiter compensates for user voice variance. The belt pack may also include a two position gain switch for normal and high-noise environments. The *remote mic-kill* function at the base station enables belt pack microphones to be shut off from another location to conveniently eliminate microphone pickup. This is done by sending a 20–24 kHz ultrasonic signal down the audio line, turning off the talk gate on each unit on the line. A visual call signal is provided by high-intensity LEDs on the belt pack to alert operators who have removed their headsets.

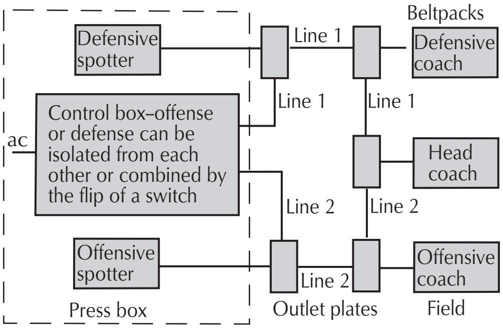


Figure 43-15. Football communication system. Courtesy Telex Communications, Inc.

43.2.7 Telephone Interface

Broadcast intercoms are often connected to a telephone line. This is accomplished through a microprocessor-based telephone interface that provides communications between a wet dial-up phone line and an intercom system. The interface is ideal for the broadcast industry, and is specifically designed to connect a telephone line to ENG and EFP trucks, production studio consoles, and TV facilities. The device automatically answers incoming calls to the intercom system, and automatic forward-nulling circuitry adjusts internal hybrids on both sides of the line to achieve a null of up to 40 dB in less than 1⁄10 s. The devices also include automatic gain control to insure that the incoming telephone audio remains at a constant level. It is also capable of automatic dial-up IFB, enabling a field crew to directly access a preset IFB circuit to immediately communicate with the studio crew.

When the device is set up for automatic answering, an incoming call is automatically answered, and the ringing is indicated on the master station and can illuminate a light at all intercom stations. The interface can automatically hang up or release the line if it detects a dial tone, resulting from either an intentional hang-up or a disconnection due to a line problem. If an audio program becomes too loud or distracting, it can be momentarily or permanently interrupted by any local intercom user.

Because telephone-to-intercom interfaces are used with standard telephone lines, frequency response is limited to 250 Hz–3.4 kHz ±3 dB. Automatic volume control (AVC) is about 20 dB and the depth of the null is greater than 30 dB, 200 Hz–8 kHz.

Many of the digital matrix intercoms offer extensive direct dial-in access to the system from any touch-tone telephone in the world. They may include up to 50 two digit DTMF codes that can be used to select any station, group, program source, or IFB circuit in the system.

43.2.8 Headsets

Most headsets are designed to work with all major types of communication systems including party line and matrix intercoms. They are used in intercom and sportscaster/announcer applications where audio quality, reliability, comfort, and the ability to hear and talk in noisy environments are of prime importance.

Light weight and comfort as well as durability are of major importance. Many headsets are made of flexible composite materials, which will not be damaged if dropped, thrown, or stepped on.

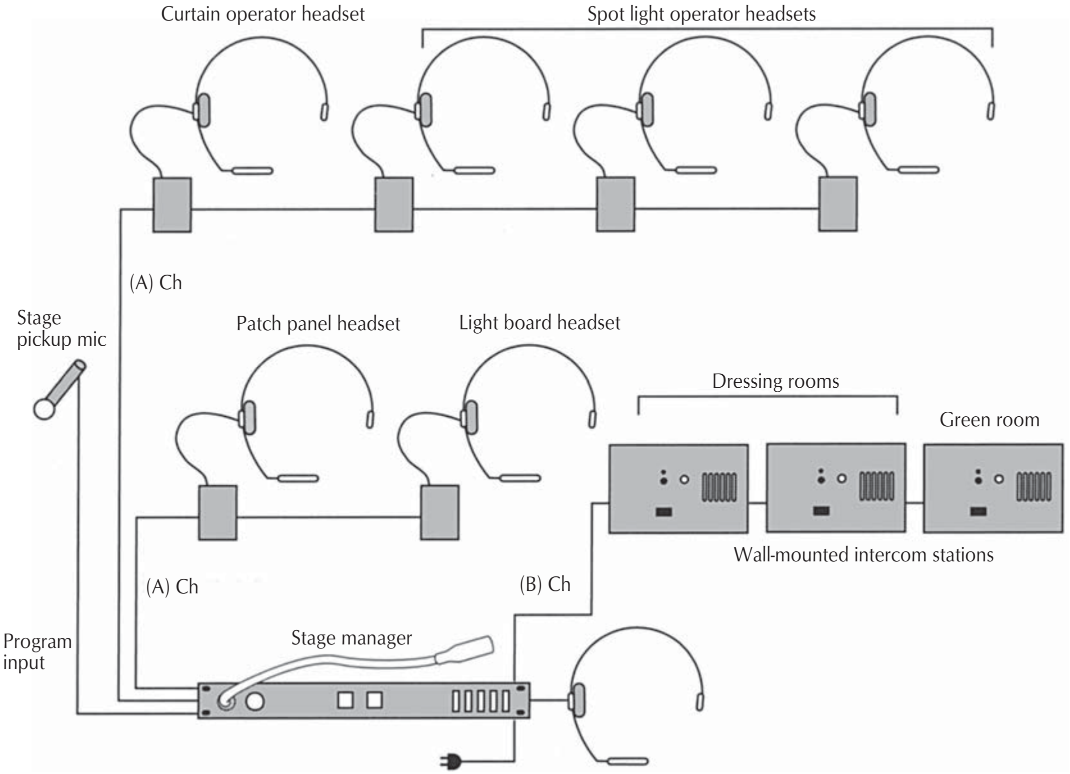

Curtain operator headset
Spot light operator headsets
(A) Ch
Stage pickup mic
Patch panel headset
Light board headset
Dressing rooms
Green room
Wall-mounted intercom stations
(A) Ch
(B) Ch
Program input
Stage manager

Figure 43-16. Theater intercom system. Courtesy Clear-Com Communication Systems.

Attention must be paid to acoustical and electrical isolation between the microphone and earphone(s) to minimize the common problem of crosstalk in multiple channel intercom systems. Also by using dual chamber foam filled ear cushions, the acoustical isolation of earphones allows for comfort and low ear fatigue in high-noise environments.

Broadcast-quality microphones require noise rejecting abilities with wide-frequency response and good resistance to breath and wind noise as the microphone is usually very close to the lips. Boom microphones can be adjusted to any position, and located as a right or left hand headset with positive detent stops. In addition, the boom can be bent into any required position. Often swinging the boom up shuts off the microphone to eliminate feedback and unwanted noise.

The headset cable is specially designed to minimize crosstalk between the microphone and the earphone. The wire stranding is a special composition for flexibility and resistance to breakage.

Earphones have specially contoured wide-bandwidth frequency response and a sensitivity of 94 dB SPL with 1 mW of power. This reduces amplifier power, increasing battery life.

Chapter 44

The Fundamentals of Display Technologies

by Alan Brawn

44.1 The Effect of Display Specification 1579
  44.1.1 Lumens 1579
  44.1.2 Candelas per Meter Squared (cd/m2), or Nits 1579
  44.1.3 Contrast 1579
44.2 Display Color 1580
44.3 Display Resolution 1580
44.4 Display Scanning 1580
44.5 Aspect Ratios and Screen Formats 1581
  44.5.1 Common Aspect Ratios 1581
44.6 Scaling 1581
  44.6.1 Analog Image Display 1581
  44.6.2 Digital Fixed Matrix Display Scan Conversion 1582
44.7 Video Signals 1582
  44.7.1 What Comprises a Video Signal? 1582
  44.7.2 Composite (aka NTSC) 1582
  44.7.3 Y/C (aka S-Video) 1582
  44.7.4 Component Video 1582
  44.7.5 VGA (Video Graphics Array) 1582
  44.7.6 DVI (Digital Video Interface) 1582
    44.7.6.1 HDMI (High-Definition Multimedia Interface) 1583
44.8 Digital Display Technologies 1583
  44.8.1 Plasma Display Technology 1583
    44.8.1.1 PDP Characteristics 1583
    44.8.1.2 PDP Operates in the Following Manner 1583
  44.8.2 Liquid Crystal Displays 1584
    44.8.2.1 LCD Characteristics 1584
    44.8.2.2 LCD Operates in the Following Manner 1584
  44.8.3 Digital Light Processing 1585
    44.8.3.1 DLP Characteristics 1585
    44.8.3.2 DLP Operates in the Following Manner 1585
    44.8.3.3 New DLP BrilliantColor™ Color Wheel Technology 1586
  44.8.4 Liquid Crystal on Silicon 1587
    44.8.4.1 LCoS Characteristics 1587
    44.8.4.2 LCoS Operates in the Following Manner 1587
  44.8.5 Organic Light Emitting Diode 1587
    44.8.5.1 OLED Characteristics 1587

44.8.5.2 OLED Works in the Following Manner 1588
44.8.6 Light Emitting Diode 1588
44.8.6.1 LED Characteristics 1588
44.8.6.2 LED Operates in the Following Manner 1588
44.9 Resolution 1589
44.10 Conclusion 1589

In this chapter we will explore the fundamentals of display technologies and how each technology does its work. We will begin by covering display specifications, video and computer signals, and finally the display technologies themselves providing the full context of how an image is produced on screen.

44.1 The Effect of Display Specification

In most audiovisual systems design, the display is the key focal point in the room. With this in mind, it is a requirement to match the display to the environment and the explicit needs set forth in the original sales proposal and system design. It is necessary to understand the specifications relating to display technologies, in order to properly design for individual applications. The key considerations are:

- Brightness.
- Contrast.
- Color.
- Resolution.
- Scaling.

Brightness is the element that we are most familiar with. It is the measurement of light falling on a surface, or light emitting from a source such as a plasma, LCD, or OLED flat panel display. It is most commonly stated in two units of measurement when relating to display technologies.

44.1.1 Lumens

A *lumen* is a measurement of light falling on a surface, such as a projector illuminating a screen. The measurement is taken using a light photometer pointed at the projector lens thus measuring the light output of the projector itself.

A lumen is equal to one foot-candle falling on one square foot of area.

Properly specified, it is referred to as an ANSI lumen, which infers adherence to the American National Standards Institute method of measurement utilizing a nine zone pattern of rectangles and averaging the light measurement from each of the nine zones.

Since there are no mandatory/standardized methods of verifying lumen specifications from a given display manufacturer, the actual light output may vary as much as 20% less than the published specifications. This necessitates testing each display for actual lumen light output prior to specifying a specific projector in each application.

While lumen light output is the common specification, it is really foot Lamberts or the light reflected from the screen surface to the viewer that is most important. This necessitates taking the screen surface gain and ambient light in the room into consideration when specifying a projector.

Lumen light output will range from 100 lumens in a tiny pico projector up to over 30,000 lumens in a large rental and staging or digital cinema projector.

44.1.2 Candelas per Meter Squared (cd/m²), or Nits

Candelas per meter squared, or cd/m², is a unit of measure that may also be referred to as *nits* and is typically used in the measurement of light emitting from a flat panel display such as plasma, LCD, or OLED directly at the viewer.

Broken down, candela, abbreviated as *cd*, is a term that originated in the days when candles were used in theaters.

For our purposes, candelas per meter squared measures the light properties radiating from a one-meter-square surface, providing a technical frame of reference for the performance of a display's black level, peak brightness, grayscale, and gamma readings.

Candelas per meter squared are more accurate than lumen light output measurements and can be measured using the same nine zone ANSI pattern. Since the screen reflectivity or gain is taken out of the equation it is less complex than a projector and screen combination.

Typical candela per meter squared measurements vary from 300 cd/m^2 for a 19 inch desktop monitor to the latest LCD displays providing 1500 cd/m^2 in sizes up to 108 inches diagonal.

It is generally agreed that contrast is the element in a picture that provides the appearance of quality in an image. Poor contrast makes the image appear washed out while good contrast gives us excellent depth of field and much more detail in a picture. It also gives us the appearance of higher resolution while not actually providing more lines of resolution or more pixel density in the image.

44.1.3 Contrast

Contrast is the range of light and dark values in a picture.

- It is stated as the ratio between maximum and the minimum brightness values—e.g., 1000:1.
- Low contrast is shown mainly in shades of gray.

- High contrast is shown as blacks and whites with very little gray.
- In digital technologies, contrast is the difference between the luminance or brightness of an ON pixel and an OFF pixel.

Contrast is used as a marketing specification and is the most misstated of all specifications from manufacturers. The proper method to measure contrast is by using a sixteen zone black and white ANSI test pattern and comparing the average of the dark rectangles to the white rectangles to get the proper contrast ratio. As a point of reference, when measured in this manner, the most expensive digital cinema projectors in the world will produce approximately 500:1 contrast and a typical boardroom projector will produce less that 100:1 contrast in a typical lighting condition.

44.2 Display Color

Each display technology produces the color we see in a different manner but they all utilize the primary colors of red, green, and blue as well as the secondary colors of cyan, magenta, and yellow to create the full color spectrum we see onscreen.

Going back to our physics class in high school, we remember that white light, when viewed through a prism, produces a veritable rainbow of colors, known as the full color or electromagnetic spectrum. The easiest way to remember the visible color spectrum is the name ROY G BIV (red, orange, yellow, green, blue, indigo, and violet).

Infrared and ultraviolet light are not visible, and fall at the far extremes of the spectrum.

In the world of professional audiovisual, you may encounter what is known as the CIE chromaticity diagram (color chart). This chart illustrates the full color spectrum, including wavelengths of light measured in nanometers and color temperature, measured in degrees Kelvin. As a specific point of reference, you can see the overlap of red, green, and blue, producing white light.

In modern display technologies, color is created in a variety of ways, by the manipulation of white light:

- Color is produced from the light of a projector lamp shining through a color wheel, in the case of single chip DLP (digital light processing).
- Color is produced from the light of a projector lamp shining through a transmissive display device, or bouncing off a reflective display device, such as LCD, three chip DLP, or LCOS (liquid crystal on silicon).
- Color is produced from an emissive (light producing) display device, such as a CRT, plasma display, or OLED.
- Color is produced from the backlight of a transmissive display device, such as a flat panel LCD.

Color and color space can be calibrated in the majority of displays by using instruments known as colorimeters. These devices measure the visible color spectrum and permit the technician to calibrate for a specific color temperature depending on the application. Most displays are set for 6500 Kelvin, also known as D65, which replicates an image in a full daylight mode.

44.3 Display Resolution

The resolution of digital display technologies is fixed as in the reference to fixed matrix displays. Resolution relates directly to visual acuity and what the eye can see. Each display technology differs in the spaces between the pixels and this is called the fill factor. The displays with the highest fill factor appear to have less of what is known as the screen door effect thereby providing a look closer to that of an analog image of 35 mm color film or CRT displays. The higher the number of pixels, the more detail in an image.

- In digital displays, resolution is the number of pixels (picture elements or individual points of color) contained on a display, expressed in terms of the number of pixels on the horizontal axis and the number of pixels on the vertical axis—e.g., 1920 × 1080.
- The sharpness of the image on a display depends on the resolution and the size of the display. The same pixel resolution will be sharper on a smaller display and gradually lose sharpness on larger display because the same number of pixels is being spread out over a larger number of inches.
- In terms of fill factor, LCD has the most space between pixels, with DLP providing more fill factor, and LCoS providing the highest fill factor available today.

44.4 Display Scanning

Television signals and compatible displays are typically interlaced, and computer signals and compatible displays are typically progressive (noninterlaced). These two formats are incompatible with each other; one would need to be converted to the other before any common processing could be done.

Designed for analog NTSC television, interlaced scanning is where each picture, referred to as a frame, is made up of two separate subpictures, referred to as fields, so two fields make up a frame. An interlaced picture is drawn on the screen in two passes, by first scanning the horizontal lines of the first field and then retracing to the top of the screen and scanning the horizontal lines for the second field in between the first set. Field 1 consists of lines 1 through 262½, and field 2 consists of lines 262½ through 525. A television scans 60 fields every second (thirty odd and thirty even). These two sets of thirty fields are combined to create a full frame every 1⁄30th of a second, resulting in a display of thirty frames per second. Drawbacks to interlaced scanning compared to progressive scanning include lower resolution, flicker, aliasing, and image artifact quality issues.

Progressive scan differs from interlaced scan in that each line (or row of pixels) in the signal is drawn in a sequential order rather than an alternate order, as is done with interlaced scan. In short, with progressive scan, the image lines (or pixel rows) are scanned in numerical order (1, 2, 3) down the screen from top to bottom, instead of in an alternate order as done in interlaced scanning. By progressively scanning the image onto a screen every 60th of a second rather than "interlacing" alternate lines every 30th of a second, a smoother, more detailed, image can be produced on the screen. The benefit is the viewing fine details, such as text, and is also less susceptible to interlace flicker and basically eliminates aliasing on the edges of objects in a picture. The drawback to progressive scan is that it requires more bandwidth to display the images onscreen.

44.5 Aspect Ratios and Screen Formats

Aspect ratio refers to the shape of the images we see on screen, but just what comprises an aspect ratio? Aspect ratio is typically described as the ratio of screen width to screen height. There are two common aspect ratios. The first is that of a standard televisio, which has a 4:3 (referred to as 4 by 3) aspect ratio. Also note that the television aspect ratio is listed as 1.33:1. This is another way of listing aspect ratios—dividing the width by the height (e.g., 4⁄3 = 1.33). This is referred to as 1.33:1 or 1.33 to 1. A widescreen display, such as a plasma panel, will usually have a 16 by 9 aspect ratio (16:9). Since 16/9 = 1.78, the aspect ratio is also known as 1.78:1 or 1.78 to 1.

44.5.1 Common Aspect Ratios

- **4 × 3 (1.33:1).** This is the standard television format used throughout the second half of the 20th century. This is both typical computer and NTSC broadcast video.

  Note: 1280 × 1024 is actually 5:4 aspect, not 4:3.
- **16 × 9 (1.78:1).** This is the common format for widescreen DVD movies, HDTV (720p and 1080i) and widescreen computer resolutions (1280 × 720, 1920 × 1080, etc.).
- **13 × 7 (1.85:1).** This is the standard aspect ratio for theatrical release film prints.
- **29 × 9 (Cinemascope–2.35:1).** A very wide screen format used for theatrical release movies, and some new DVDs.

44.6 Scaling

In the realm of digital display technologies, there is quite often a mismatch between the resolutions of the display itself and the signals or sources coming into the display. This mismatch necessitates the incorporation of a process known as scaling or scan conversion.

By definition, a digital display can also be referred to as a fixed matrix display, with a finite number of horizontal and vertical pixels—e.g., 1024 × 768.

In many instances the actual resolution of the input signals and the physical resolution of the display do not match. The mismatch requires what is known as scaling.

While a scaler can be an outboard device, in most instances today, it is built into the display device.

Scaling, which is sometimes called scan conversion, refers to a process of taking a higher-resolution signal, and modifying it to be displayed on a lower-resolution device, or a lower-resolution signal, and modifying it to be displayed on a higher-resolution device.

44.6.1 Analog Image Display

In an analog display, such as a CRT, scan conversion is not required, because the output of the display is infinitely adjustable to match the input signal entering the display.

The pixels, more properly referred to as rare earth phosphor spots, are adjusted in width and position on the CRT to match the source image.

44.6.2 Digital Fixed Matrix Display Scan Conversion

In a digital or fixed matrix display, where the pixels are in a fixed size and position, the input signal may or may not match. Scan conversion is a necessary process to fit the analog source image across multiple pixels in the display. Depending on the display technology, 20–30% of image information may be lost.

For example, if we have a fixed matrix display at 1024 × 768 and want to input a signal that is 800 × 600, a mathematical algorithm is employed where the signal information in the lower-resolution signal is reduced to a mathematical equation and fit into the higher-resolution display. The same process takes place where a higher-resolution image is fit mathematically into a lower-resolution display.

In all instances where scaling and scan conversion take place there is lost information. The quality of the scaler or scan converter varies with each display and the most fidelity in an image takes place where the signal and the display resolutions match each other.

44.7 Video Signals

In order to further the understanding of displays and display technologies, it is necessary to gain a basic comprehension of what comprises the different types of video signals in use today. We will now examine the core components of all signals, and their transmission standards.

44.7.1 What Comprises a Video Signal?

Chrominance noted as (C) is the hue or color with saturation in the red, green, and blue channels of a signal.

Luminance noted as (Y) is the amount of light in each red, green, and blue channel.

Without the chrominance in a signal, the picture is black and white.

44.7.2 Composite (aka NTSC)

An analog composite video signal is used in most home applications. It combines the chrominance and luminance, along with a sync signal into one cable. This facilitates the broadcast of the NTSC television signal to our homes.

44.7.3 Y/C (aka S-Video)

This is still a composite signal, but one that nearly separates luminance and chrominance to provide a more precise color reproduction on the screen.

44.7.4 Component Video

Commonly known as RGB (RGB sync, RGB with H and V sync, RGB sync on green).

This type of signal totally separates red, green, blue, and sync to give clearer definition to each of the color channels.

Component is never used for broadcast due to its excessive bandwidth requirement in the green channel. Note that the sync signal can be H/V (horizontal and vertical) or sync on green.

One version of a component signal is commonly known as YPbPr. In the broadcast community, this is known as a color difference signal. Since a normal RGB signal requires too much bandwidth to broadcast the dominant green channel, YPbPr, as a component signal, extrapolates the green signal by subtracting from the luminance channel (Y) both the blue component (Pb) and the red component (Pr), leaving the green component.

This allows the economical broadcast of a component signal by reducing the bandwidth needed by eliminating the dedicated green signal.

44.7.5 VGA (Video Graphics Array)

VGA is the analog display standard for the PC. VGA uses an analog monitor, and PC display adapters to output analog signals. All PC CRTs and most flat panel monitors accept VGA signals, although newer flat panels may also have a DVI interface for display adapters that output digital signals.

VGA may refer to the physical 15-pin VGA socket on a PC in order to contrast it with a digital DVI socket for flat panels. Or, VGA may also refer only to the original VGA resolution of 640 × 480 and 16 colors.

44.7.6 DVI (Digital Video Interface)

DVI is a multipin connection used for passing standard-definition and high-definition digital video signals, found on HDTV tuners, a growing number of DVD players, HDTV-ready televisions, and some computer displays. DVI connections transfer video signals in pure digital form, which is especially beneficial if you're using a fixed-pixel display (like a LCoS, plasma, LCD, or DLP TV). Signals are encrypted with HDCP high-bandwidth digital content protection) to prevent content from being re-recorded and pirated.

There are different kinds of DVI connections. DVI-D, which is the type of DVI connection found on most home video gear, carries digital-only signals. DVI-I, used with some computer video cards, is capable of passing both digital and analog video signals. Some TVs feature DVI-I inputs for greater hookup flexibility.

44.7.6.1 HDMI (High-Definition Multimedia Interface)

HDMI is the second generation digital interface that evolved out of the DVI standard.

HDMI is a multipin connection used for passing standard- and high-definition digital video signals, as well as multichannel digital audio, through a single cable. These connections are usually found on newer HDTV tuners, and a growing number of DVD players, HDTV-ready televisions, and home theater receivers. HDMI cable accommodates up to 5 Gbps bandwidth, so it can simultaneously transfer pure digital video and audio signals without compression (even HDTV video).

HDMI works especially well with a fixed-pixel display (like a LCoS, plasma, LCD, or DLP TV), and is backwardscompatible with most DVI connections. Signals are encrypted with HDCP (high-bandwidth digital content protection) to prevent recording.

Although many first generation HDMI-equipped components only pass two-channel audio signals, HDMI can carry up to eight discrete audio channels, making it forward-compatible with 7.1 sound systems. That means you can pass digital video and multichannel audio signals between newer HDMI-equipped components along a single cable.

44.8 Digital Display Technologies

In the early days of the audiovisual industry, it was necessary to immerse oneself in the tiniest details of technology and how it operated. In today's market, it is necessary to understand the basic function of various technologies, and more specifically, how the basic functions affect the final design, and the solutions presented to the client and for the specific project.

We will examine the characteristics and basic functions and operation of the following:

- PDP (plasma).
- DLP (digital light processing).
- LCD (liquid crystal display).
- LCOS (liquid crystal on silicon).
- OLED (organic light emitting diode).
- LED (light emitting diode).

44.8.1 Plasma Display Technology

Of all fixed matrix display technologies, plasma, or PDP displays most closely replicate the smooth image from a 35 mm film projector and a CRT. Plasma displays are emissive in nature, and utilize a similar rare earth phosphor to a CRT to provide color saturation for the display, Fig. 44-1.

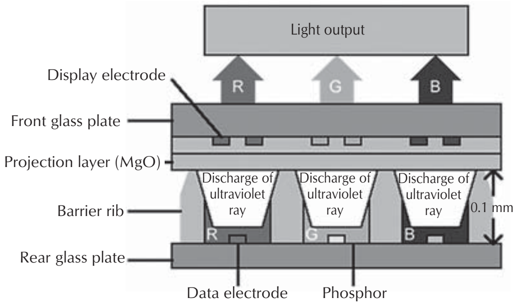


Figure 44-1. Plasma monitor.

44.8.1.1 PDP Characteristics

- 3 to 4 inch thick displays (wall or base mount).
- 60 to 500 pounds.
- Panel sizes 37 inch, 40 inch, 42 inch, 43 inch, 46 inch, 50 inch, 55 inch, 60 inch, 61 inch, 63 inch, 71 inch, 103 inch, and 150 inch.
- 16:9 aspect ratio panels.
- PDP combines the pixel structure of LCD with the color generation of a CRT.
- No radiation or high voltage emissions.
- Fast response time.
- High contrast.
- Deep color saturation.

44.8.1.2 PDP Operates in the Following Manner

- The cells are filled with a xenon and neon gas mixture.
- A controlled current is passed through the gas.
- Ultraviolet rays are produced by the current energizing the gas, creating a plasma.
- Ultraviolet rays hit the red, green, and blue phosphors applied inside the cells.
- Visible light is produced by the ultraviolet rays exciting the rare earth phosphors.
- Voltage is applied to one of three terminals on a pixel. The voltage discharges through the pixel to a second electrode ionizing a rare gas (creating a plasma) in the process. The ionization creates UV light, which excites an R,G, B phosphor causing it to glow (like a CRT). Brightness variation is achieved by controlling the number of pulses of light that our eyes integrate to produce impression of dim or bright areas.

44.8.2 Liquid Crystal Displays

Liquid crystal displays have become ubiquitous. As the foundation for modern computer and cell phone displays, LCD technology is used for large flat panel displays as well as three chip LCD projectors. No matter the application, LCD technology and how it works is similar in the way it fundamentally operates.

44.8.2.1 LCD Characteristics

- 3 to 4 inch thick displays (wall or base mount).
- 60 to 400 pounds.
- Panel sizes ranging from under 8 to 108 inch panels.
- 4:3, 16:9, and 16:10 aspect ratio panels.
- No radiation or high-voltage emissions.
- Low power consumption.
- High resolution, up to 4 × HDTV.
- Ideal for computer display and digital signage.

44.8.2.2 LCD Operates in the Following Manner

There's far more to building an LCD than simply creating a sheet of liquid crystals. The combination of four facts makes LCDs possible:

- Light can be polarized.
- Liquid crystal can transmit polarized light or change the plane of polarization.
- The structure of liquid crystals can be changed by electric field.
- There are transparent substances that can conduct electricity.

To create an LCD, you take two pieces of glass with polarizing films applied.

A polyimide film is applied to the liquid crystal side of the glass and then mechanically rubbed to produce microgrooves.

The two glass plates are assembled together with a carefully controlled gap dimension.

When LC material is introduced to this cell, the layers adjacent to the polyimide will align with the microgroove directions resulting in a helical structure of LC molecules between the two glass plates, Fig. 44-1.

Liquid crystal displays come in two basic configurations: flat panel displays and projection displays. Both variations utilize the same basic LCD principle, but differ in the way that they are illuminated.

In the flat panel, or desktop display, the illumination comes from bright cold cathode fluorescent lights behind the display.

In projection LCD displays, the illumination comes from a bright lamp reflecting off of the LCD and onto the screen.

LCD monitors make use of thin film transistors (TFT). TFTs are small switch transistors and capacitors that sit on a glass substrate in the LCD structure, Fig. 44-2.

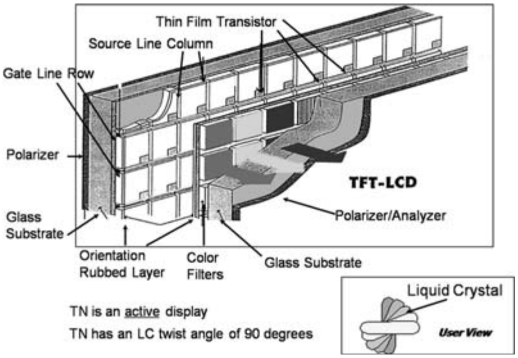


Figure 44-2. TFT-LCD technology.

Each pixel is controlled by one up to four of these TFTs. To ignite a particular pixel, power is applied to the correct column and row (just like passive matrix).

Any pixels on the same row and column that are not targeted simply pass the current on. The transistor at the target pixel stops the current. The capacitor takes the current and stores it. It is then able to hold that charge until the next screen refresh.

Also, by adjusting the amount of voltage to each pixel, you can control the amount that the crystals will untwist, thereby allowing varying degrees of color.

For an LCD monitor to produce color, each pixel on the screen has to have three subpixels, each being a primary color (red, blue, and green). In this aspect, color LCDs work the same way as the color CRT. By taking each of the three colors, each having 256 possible shades, and blending it all together, the color active matrix LCD has a possible palette of 16.8 million colors. Each subpixel has a transistor/capacitor and with this design process, one can see that there are millions of transistors necessary to formulate a full TFT screen.

In an LCD monitor, the light source is behind the panel and illuminates the display from behind. Typically the lighting is a florescent type but the most recent development in illuminations is via side emitting LED display, which improve uniformity, durability, brightness, and the life of the backlight.

LCD projectors utilize three LCD panels or chips as the imaging devices but unlike LCD monitors, they differ in the way color and illumination are derived. By looking at the LCD light path illustration, Figs. 44-3, we begin with a metal halide lamp for illumination. The lamp approximates pure white light from which the colors of the spectrum can be extrapolated. Color (RGB) is achieved by incorporating dichroic mirrors or filters into the light path. The dichroic mirrors filter out all of the unwanted color spectrum and pass on a narrow band of color coordinates of red, green, and blue, permitting each of those colors in a pure form to be transferred to the main optical prism or combiner just behind the projection lens.

LCD projectors come in various sizes, shapes, light outputs, and resolutions.

Typical native resolutions for commercial LCD projectors are 800 × 600, 1024 × 768, and 1280 × 1024.

In terms of weight, the brighter the projector, the bigger the lamp housing requirement and hence the heavier the projector.

Modern LCD projectors vary in size from 5 lbs to over 50 pounds for the high-brightness models.

Brightness has long been the holy grail of projectors and LCD with flat panel brightness reaching 1500 cd/m² and projector brightness achieving 15K lumens.

44.8.3 Digital Light Processing

Digital light processing was developed by Dr. Larry Hornbeck of Texas Instruments and brought to market in the mid-1990s. It is fundamentally a digital light switch that is used in projection applications as far reaching as tiny pico projectors to be inserted in cell phones all the way to digital cinema projectors replacing 35 mm film in movie theaters. Its compact size along with single chip and three chip variations make it unique in the world of display technologies, Fig. 44-4.

44.8.3.1 DLP Characteristics

- Projection technology, no fixed screen size.
- Single or three chip configurations.
- 16:9 and 16:10 aspect ratio panels.
- No radiation or high-voltage emissions.
- Low power consumption.
- High resolution, up to 2 K.
- High brightness and contrast.
- Does not require polarized light.

44.8.3.2 DLP Operates in the Following Manner

- DLP™ is based on an optical semiconductor called a digital micromirror device, or DMD.
- The DMD is an extremely precise light switch that enables light to be modulated digitally via millions of microscopic mirrors arranged in a rectangular array.
- Each mirror is spaced less than 1 micron apart.
- These mirrors are literally capable of switching on and off thousands of times per second and are used to direct light toward, and away from, a dedicated pixel space.
- When the display is off, all of the mirrors are flat.

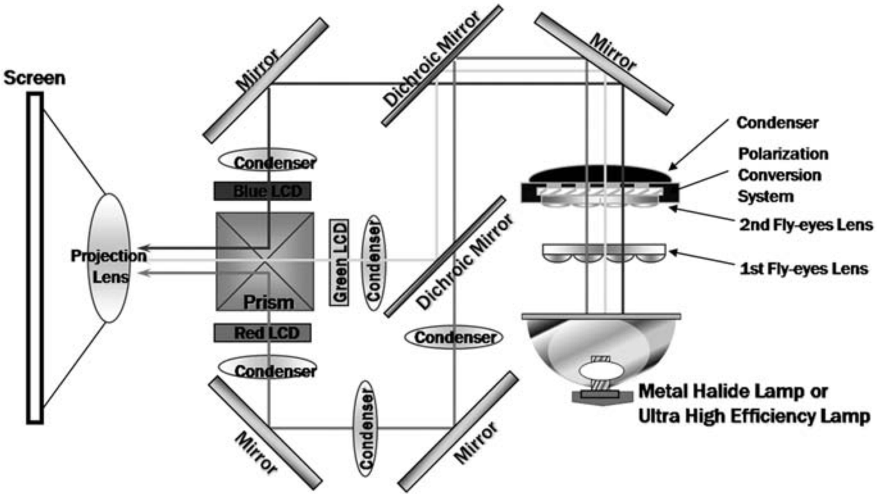

Screen
Mirror
Dichroic Mirror
Mirror
Condenser
Polarization Conversion System
2nd Fly-eyes Lens
1st Fly-eyes Lens
Condenser
Blue LCD
Projection Lens
Prism
Green LCD
Condenser
Dichroic Mirror
Red LCD
Condenser
Condenser
Metal Halide Lamp or Ultra High Efficiency Lamp
Mirror
Condenser
Mirror

Figure 44-3. LCD projection TV optical path.

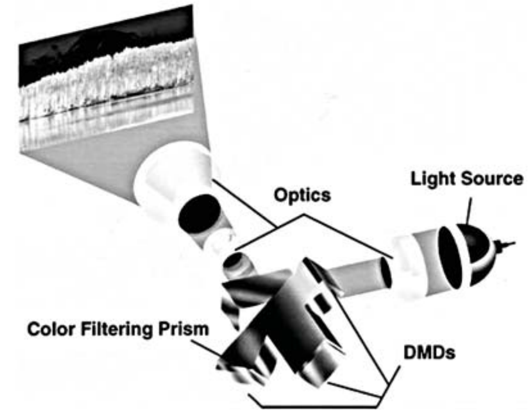


Figure 44-4. Three chip DLP projection TV.

- When the display is turned on and the chip begins transmitting the signal, the mirrors flip back and forth thousands of times per second.
- Mirrors in the on position reflect the light through a projection lens and onto the screen. The longer a mirror is in the on position, the lighter the pixel it creates. Mirrors that are off for longer periods create darker pixels, and mirrors that are always off create black pixels. By varying the length of time that the mirrors point toward the projection lens, the DMD creates up to 1024 shades of gray.
- The gray pixels combine on the screen to create a progressive, fully digital monochrome image.
- To add color to the picture, the single chip DLP system uses a color wheel, Fig.44-5.
- The color wheel is a transparent, spinning wheel with red, green, and blue. The light passing through each section turns red, gree, or blue.
- The system's processor synchronizes the spinning of the wheel with the action of the mirrors. Together, the DMD and the color wheel can create 256 shades of each primary color.
- Each pixel of light on the screen is red, gree, or blue at any given moment. The colors are then blended to create the desired colors of the image.
- With DLP projectors, a small number of people might experience a rainbow effect when watching a DLP projection, especially when they change their focus from one part of the image to another, seeing the individual component colors.
- This happens only in DLP systems that use a segmented color wheel, not in systems that use one DLP chip for each primary color.

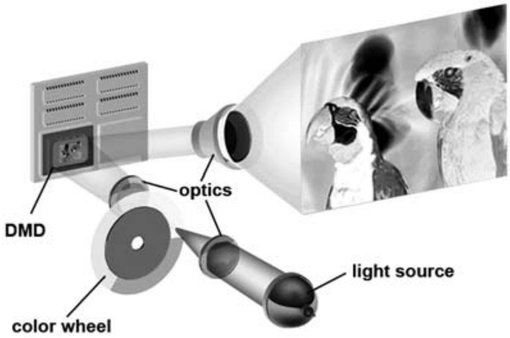


Figure 44-5. Single chip DLP technology.

- A number of home theater systems use color wheels with additional segments, two segments of each color, or sequential color recapture (primary colors arranged in a spiral instead of in segments) in order to reduce the appearance of the rainbow effect.
- New BrilliantColor™ color wheel technology reduces the appearance of the rainbow effect. In the newer DMD generations, a light-eating dark metal coat is applied to the interior of each chip, preventing stray light from traveling to the screen when mirrors are switched off. This improvement increases contrast ratios from 1200:1 to >2000:1 and higher.
- Increased mirror tilt angle (from ±10 to ±12°), brings 20% more light to the screen for greater brightness.
- Double data rate technology allows a DMD chip to tilt toward or away from its light source twice as fast, allowing more accurate grayscale reproduction.

44.8.3.3 New DLP BrilliantColor™ Color Wheel Technology

Historically, most display devices would render a scene using the three primary colors, red, green, and blue.

This limits available colors that can be displayed, making it difficult to display brilliant yellows, magentas, and cyans that are commonly found in natural scenes.

BrilliantColor™ technology adds yellow, cyan, and magenta colors to the color wheel, maintaining bright whites while providing deeper red, green, and blue colors.

BrilliantColor™ provides brightness increases in nonprimary colors and boosts overall color intensity.

BrilliantColor™ provides flexibility in color wheel design allowing for bright, large color gamuts and differentiation from OEMs.

44.8.4 Liquid Crystal on Silicon

Liquid crystal on silicon combines the best of both worlds of LCD and DLP. It is a reflective technology like DLP but uses liquid crystals instead of moving mirrors to control the light transmission levels of the individual pixels. The benefit of LCoS is that it has excellent color and contrast capabilities and has the highest fill factor of any current digital display. It is also capable of 4 K resolution and is a competitor with DLP for digital cinema applications, Fig. 44-6.

44.8.4.1 LCoS Characteristics

- Projection technology, no fixed screen size.
- Three chip configuration.
- 16:1 aspect ratio panels.
- No radiation or high voltage emissions.
- Low power consumption.
- High resolution, up to 4 K.
- High brightness and contrast.
- High fill factor.

44.8.4.2 LCoS Operates in the Following Manner

- LCoS technology is a reflective liquid crystal modulator where electronic signals are directly addressed to the device.
- The LCoS device has an X-Y matrix of pixels configured on a CMOS single crystal silicon substrate mounted behind the liquid crystal layer using a planar process that is standard in IC technology.
- The liquid crystal is placed on top of the CMOS substrate on an array of aluminum mirrors that define each pixel.
- A glass counter electrode covers the liquid crystal to complete the structure.
- A voltage is applied to a selected pixel of the matrix in accordance with the input signal, making the liquid crystal change birefringence, thus changing the polarization direction of the incident projection light.
- The nonactive area between the pixel mirrors is minimal, only serving to separate each pixel; the rest of the electrode is active as a reflective surface, thereby providing a high aperture ratio.
- Although having the highest overall performance of any current projector technology it is held back by manufacturing yield issues and cost of components that impede its progress.

44.8.5 Organic Light Emitting Diode

Organic light emitting diode (OLED) is the newest display technology and a direct competitor for other flat panel displays such as LCD and plasma. The most obvious benefit is the nearly paper thinness of the technology. Since it is an emissive technology that does not require separate lighting it can be manufactured to create a display the thickness of a credit card. It can be made transparent and even flexible. It also has advantages in the area of low power consumption and excellent picture performance dynamics. The big issues facing OLED are manufacturing costs, and panel life, both of which are in the process of being addressed.

44.8.5.1 OLED Characteristics

- Thinnest and lightest display technology.
- Fast response time.
- High brightness.
- Low power consumption.
- Can be made transparent or flexible.

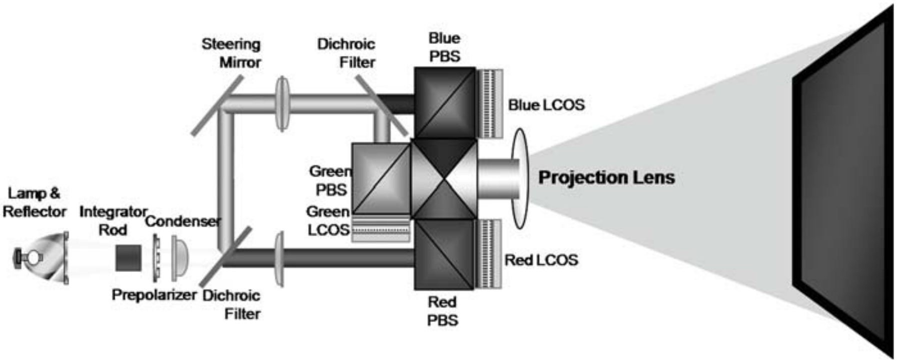


Figure 44-6. LCoS projection system.

44.8.5.2 OLED Works in the Following Manner

- The basic OLED cell structure consists of a stack of thin organic layers sandwiched between a transparent anode and a metallic cathode.
- The organic layers comprise a hole-injection layer, a hole-transport layer, an emissive layer, and an electron-transport layer.
- When an appropriate voltage (typically a few volts) is applied to the cell, the injected positive and negative charges recombine in the emissive layer to produce light (electroluminescence).
- The structure of the organic layers and the choice of anode and cathode are designed to maximize the recombination process in the emissive layer, thus maximizing the light output from the OLED device.

OLEDs are typically fabricated on a transparent substrate on which the first electrode (usually indium-tin-oxide which is both transparent and conductive) is first deposited.

Then one or more organic layers are coated by either thermal evaporation in the case of small organic dye molecules, or spin coating of polymers. In addition to the luminescent material itself, other organic layers may be used to enhance injection and transport of electrons and/or holes.

The total thickness of the organic layers is of order 100 nm.

Lastly, the metal cathode (such as magnesium-silver alloy, lithium-aluminum, or calcium) is evaporated on top.

The two electrodes add perhaps 200 nm more to the total thickness of the device. Therefore the overall thickness (and weight) of the structure is mostly due to the substrate itself.

OLEDs can be manufactured in several different types, classified by the size of molecule they use, and the type of substrate they are manufactured on. Some examples are:

- TOLED—Transparent OLED. This is manufactured on a clear substrate suitable for applications such as heads-up displays.
- FOLED—Flexible OLED. This type of OLED is manufactured into a sealed flexible substrate that can be curved, rolled, or bent.

44.8.6 Light Emitting Diode

Light emitting diodes are popping up everywhere due to their high light output and relatively low power consumption. They are finding uses in homes, automobiles, and of course high brightness outdoor displays. Their newest application is as backlight illumination for LCD flat panel displays and as a light source for small pico projectors utilizing DLP and LCoS chips.

44.8.6.1 LED Characteristics

- Extremely high brightness.
- Relatively low maintenance.
- Long life, >50,000 hours.
- Outdoor/indoor capability.
- Modular construction with scalable display sizes.
- 3 to 25 mm pixel pitches available.

44.8.6.2 LED Operates in the Following Manner

The phenomenon of *electroluminescence* was discovered in 1907 by Henry Joseph Round.

British experiments in the 1950s led to the first modern red LED, which appeared in the early 1960s.

By the mid-1970s LEDs could produce a pale green light. LEDs using dual chips (one in red and one in green) were able to emit yellow light.

The early 1980s brought the first generation of super bright LEDs, first in red, then yellow, and finally green, with orange-red, orange, yellow, and green appearing in the 1990s.

The first significant blue LEDs also appeared at the start of the 1990s, and hig-intensity blue and green in the mid-1990s.

The ultra bright blue chips became the basis of white LEDs, in which the light emitting chip is coated with fluorescent phosphors.

This same technique has been used to produce virtually any color of visible light and today there are LEDs on the market, which can produce previously exotic colors, such as aqua and pink.

Light emitting diodes (LEDs) are source of continuous light with a high efficiency.

- At the heart of a light emitting diode is a semiconductor chip, containing several very thin layers of material that are sequentially deposited onto a supporting substrate.
- The first semiconductor material that is deposited onto the substrate is doped with atoms containing excess electrons, and a second doped material, containing atoms having too few electrons, is then deposited onto the first semiconductor to form the diode. The region created between the doped semiconductor materials is known as the active layer.

- When a voltage is applied to the diode, holes (positive charges) and electrons (negative charges) meet in the active layer to produce light.
- The wavelength of light emitted by the diode is dependent on the chemical composition and relative energy levels of the doped semiconductor materials, and can be varied to produce a wide range of wavelengths.
- After being fabricated, the chip is mounted in a reflector cup connected to a lead frame, and is bonded with wire to the anode and cathode terminals.
- The entire assembly is then encased in a solid epoxy dome lens that enables emitted light to be focused, controlled by embedding tiny glass particles into the lens that scatter light and spread the light beam, or angled, via changing the shape of the lens, or the reflector cup.

44.9 Resolution

What is resolution? A simple definition of resolution is the degree of sharpness and clarity of a displayed image. In LED displays, resolution is determined by the matrix area and pitch.

The area, also known as the pixel matrix, corresponds to the number of pixels that make up the display area. In our industry, we express the matrix area in the number of pixels vertically by number of pixels horizontally, such as 16 × 64.

The pitch is defined as the distance between two pixels. The distance is measured from the center of one pixel to the center of the next pixel.

Pitch can also influence the pixel matrix for a given area. For example, a 16 mm pitch will give you a 5 × 7 matrix area, while a 10 mm pitch will give you a 8 × 11 matrix area for the same area.

Pitch determines the amount of empty space between the pixels.

Therefore, the smaller the pitch and larger the matrix area, the greater the resolution.

44.10 Conclusion

The one thing we can be certain of is that display technologies are constantly evolving with advances taking place in months not years. We can look forward to plasma and LCD displays becoming thinner and lighter with significant power consumption reduction while producing brighter displays with longer panel life. Environmentally unfriendly CCFL backlights in LCD displays will be replaced with LED and laser illumination will gain acceptance for projectors and RPTVs. OLED is set to take on conventional LCD and plasma displays and will become larger and more economically priced in the next few years. The only constant is change in the world of display technologies and we all benefit in the end.

This page intentionally left blank

Chapter 45

Surround Sound

by Joe Hull

45.1 The Origin of Surround Sound 1593
45.2 Surround from Optical Soundtracks 1593
45.3 Digital and 5.1 Surround 1593
45.4 Variations on the 5.1 Theme 1595
45.5 Surround Sound Comes Home 1596
45.6 Digital 5.1 in the Home 1597
45.7 More Home Theater Channels from Existing Content 1598
45.8 The Question of Playback Level 1599
45.9 What's Next for Surround Sound? 1600

This page intentionally left blank

45.1 The Origin of Surround Sound

The first commercially successful multichannel sound formats were developed in the early 1950s for the cinema. At the time, stereophonic sound, as it was called, was heavily promoted along with new wide-screen formats by a film industry feeling threatened by the rapid growth of television. Unlike the two-channel format later adopted for home use because of limitations imposed by the phonograph record, film stereo sound started out with, and continues to use, a minimum of four channels.

With such film formats as four-track CinemaScope (35 mm) and six-track Todd-AO (70 mm), multiple sound channels were recorded magnetically on stripes of oxide material applied to each release print. To play these prints, projectors were fitted with magnetic playback heads like those on a tape recorder (only much larger), and cinemas were equipped with additional amplifiers and loudspeaker systems.

From the outset, multichannel film sound featured several channels across the front, plus at least one channel played over loudspeakers towards the rear of the cinema. At first the latter was known as the effects channel, and was reserved for the occasional dramatic effect—ethereal voices in religious epics, for example. Some formats even switched this channel off by means of trigger tones when it wasn't needed because the magnetic track on the film was particularly narrow, and thus very hissy.

As time went on, sound mixers continued to experiment with the effects channel. In particular, because six-track 70 mm magnetic provided consistent signal-to-noise ratios on all channels, mixers began to use the effects channel to envelop the audience in continuous low-level ambient sounds. This expanded, more naturalistic application came to be known as surround sound, and the effects channel as the surround channel, Fig. 45-1.

45.2 Surround from Optical Soundtracks

Under the best conditions, the multichannel magnetic-stripe formats provided superb sound, way beyond anything the home listener could experience, and it was widely adopted in the 1950s. By the1970s, however, the expense of magnetic release prints, their comparatively short life compared to those with traditional optical soundtracks, and the high cost of maintaining the playback equipment led to a massive reduction in the number of magnetic releases and cinemas capable of playing them. Magnetic sound came to be reserved for only a handful of first-run engagements of big releases each year.

As a result, by the mid-1970s, most films were being released with only low fidelity, mono-optical soundtracks, a technology that had hardly improved since the late 1930s. Then, in 1975, came a new breakthrough: the introduction by Dolby Laboratories of a highly practical 35 mm multichannel optical release print format originally identified as Dolby Stereo.

In the space allotted to the conventional mono-optical soundtrack are two soundtracks that not only carry left and right information as in home stereo sound, but are also matrix encoded with a third center-screen channel and—most notably—a fourth surround channel for ambient sound and special effects. The matrix process in essence "folds" four channels down to two tracks on the film, and "unfolds" them in the theater by means of a sound processor-decoder (a common industry term for the process is 4:2:4), Fig. 45-2.

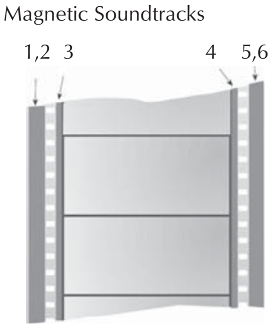
This format not only enabled multichannel sound from optical soundtracks, but higher-quality sound as well, thanks to such techniques as noise reduction and loudspeaker equalization. The result was multichannel capability on easily manufactured, compatible 35 mm optical prints that rivaled that of four-track 35 mm magnetic, which soon became obsolete.

The multichannel optical format proved so practical that within a decade of its introduction, virtually all major releases could be heard in most local cinemas in four-channel surround sound. It was dramatically improved in 1987 by the application of spectral recording (SR), a new recording process developed by Dolby Laboratories that both lowered optical track noise still further and increased headroom, making it possible to record loud sounds with wider frequency response and lower distortion.

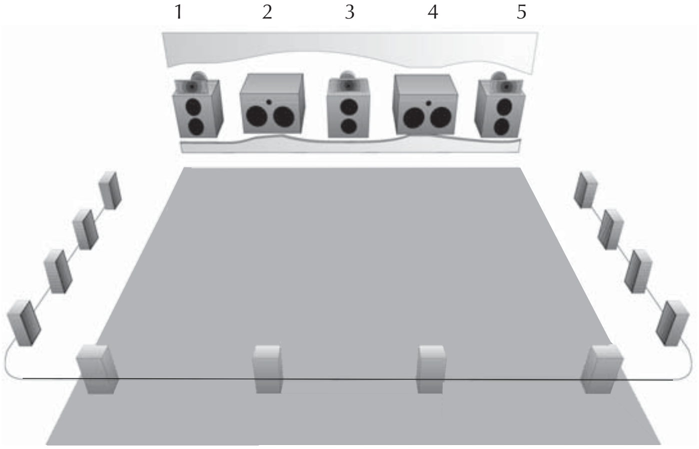
Today, virtually all 35 mm movie prints, including those with digital soundtracks, feature a matrix-encoded, four-channel SR analog optical soundtrack. The SR track makes it possible for the print to play in any theater in the world, and also acts as a backup on digital prints in case there are problems with the digital track(s).

45.3 Digital and 5.1 Surround

In the 1980s, as the success of the compact disc established with consumers the idea that digital sound meant better sound, the film industry began to investigate what it wanted by way of digital sound in the cinema. While not defining how it would be achieved, the industry agreed on several objectives, including fully discrete channels providing better channel separation than


Magnetic Soundtracks
1,2
3
4
5,6

A. 70 mm prints provided six channels recorded on four magnetic stripes.



B. Loudspeaker layout for 70 mm magnetic film.

Figure 45-1. 70 mm six-track magnetic was regarded by the film industry as the supreme audio format until the advent of multichannel digital in the early 1990s. Shown is the format as it evolved in the late 1970s, with tracks 2 and 4 carrying supplemental bass information.

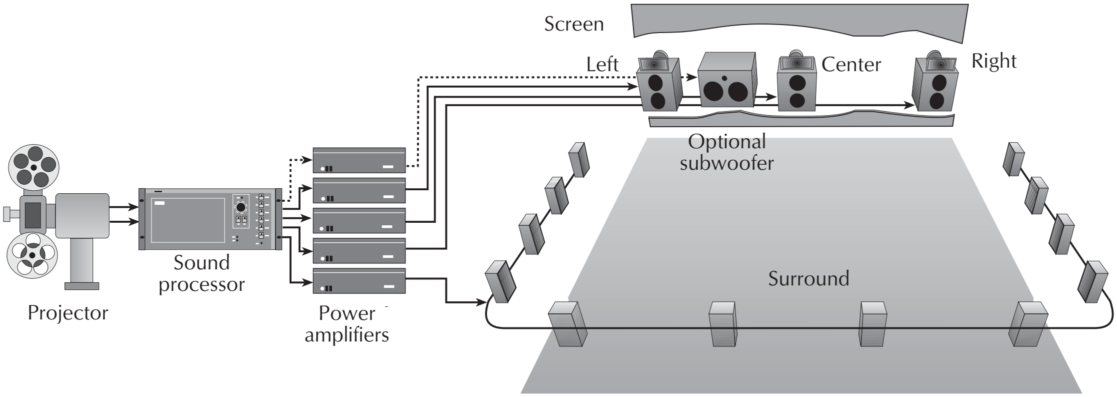

Screen
Left
Center
Right
Optional
subwoofer
Surround
Projector
Sound
processor
Power
amplifiers

Figure 45-2. Four-channel surround sound from optical soundtracks, first introduced in 1975, remains the standard analog format today, although much improved by the SR process.

matrixing, wider dynamic range, and what became known as the 5.1-channel configuration.

As shown in Fig. 45-3, the 5.1 configuration provides fully discrete left, center, right, left surround, and Right Surround channels. A sixth channel, intended for reproduction by subwoofers, carries low-frequency effects (LFE), and became known as a .1 channel because it covers only a tenth of the audible spectrum.

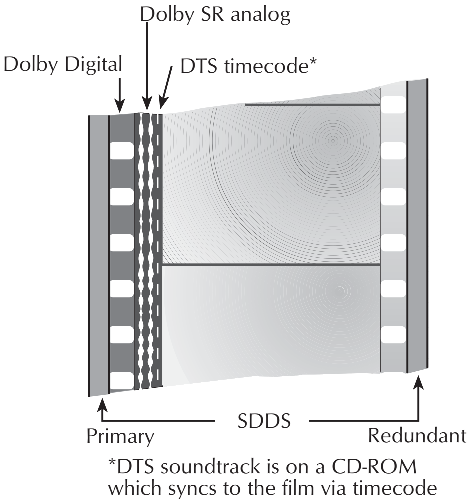

Dolby SR analog
Dolby Digital
DTS timecode*
SDDS
Primary
Redundant
*DTS soundtrack is on a CD-ROM which syncs to the film via timecode

Figure 45-3. To insure compatibility with all theaters, 35 mm quad prints feature three digital soundtracks plus an analog SR track.

5.1 has its roots in the 70 mm six-track magnetic format. Originally the six-track configuration called for five full-range screen channels and one surround channel. As the average cinema screen became smaller in the 1970s, however, the need for five full-range screen channels and elaborately panned dialogue lessened. Dolby therefore proposed that the extra two screen loudspeakers, often referred to as half-left and half-right, could be used more effectively for bass reinforcement, with the correlating tracks on the film carrying only bass information, as shown in Fig. 45-1. This technique, which was adopted by the industry and earned the nickname "baby boom," was the precursor of the .1 channel and its attendant subwoofers in today's digital formats.

The use of two surround channels was also a 70 mm contribution, beginning with some experimental 70 mm prints of the film *Superman* in 1978. What the industry came to call stereo or split surrounds was first heard by the public late in 1979 in fifteen specially equipped theaters showing 70 mm prints of *Apocalypse Now*.

The first 5.1-channel digital format was cinema digital sound (CDS), introduced in 1990 by Optical Radiation Corp. and developed in conjunction with Kodak. The CDS format placed a digital optical soundtrack on 70 mm prints in lieu of analog magnetic tracks, and experiments with 35 mm prints were underway when the venture failed. Soon thereafter, beginning in 1992, came the three competing 35 mm digital film sound formats that have survived with varying degrees of success: Dolby Digital, Digital Theater Sound (DTS), and Sony Dynamic Digital Sound (SDDS).

The three digital formats differ more in how they deliver their respective digital soundtracks than in their actual performance. Both Dolby and SDDS use optical digital soundtracks on the print, the Dolby Digital track between the sprocket holes down one side, and redundant SDDS tracks down both outer edges. DTS supplies the digital soundtrack separately on a CD-ROM disk that plays in sync with the picture by means of an optical time-code track adjacent to the analog soundtrack on the film. To insure playback in any theater, many release prints provide for all three digital formats plus analog playback, giving rise to the nickname quad print, Fig. 45-4.

45.4 Variations on the 5.1 Theme

While the 5.1 configuration became the de facto standard for multichannel digital film sound in the cinema (and for the home as well, as will be seen), both Dolby and SDDS offered producers the option of using additional channels. With SDDS, it is possible to mix for a total of seven main channels, bringing back the full-range half-left and half-right screen channels of the original 70 mm magnetic format. Thus far relatively few films have been so mixed, and relatively few theaters are equipped with the extra full-range screen loudspeakers required.

The Dolby option, called Dolby Digital Surround EX, was co-developed with Lucasfilm THX, and has achieved some success since its introduction in May of 1999 with the release of *Star Wars: Episode I–The Phantom Menace*. Surround EX offers the option of a third surround channel intended for reproduction by rear-wall surround loudspeakers, while the left and right surround channels are reproduced by the side-wall surrounds.

Not a discrete track, which is why the format's co-developers originally avoided the term 6.1, the extra surround information is matrix-encoded onto the left and right surround channels of otherwise standard. 5.1 soundtracks. This insures print compatibility with

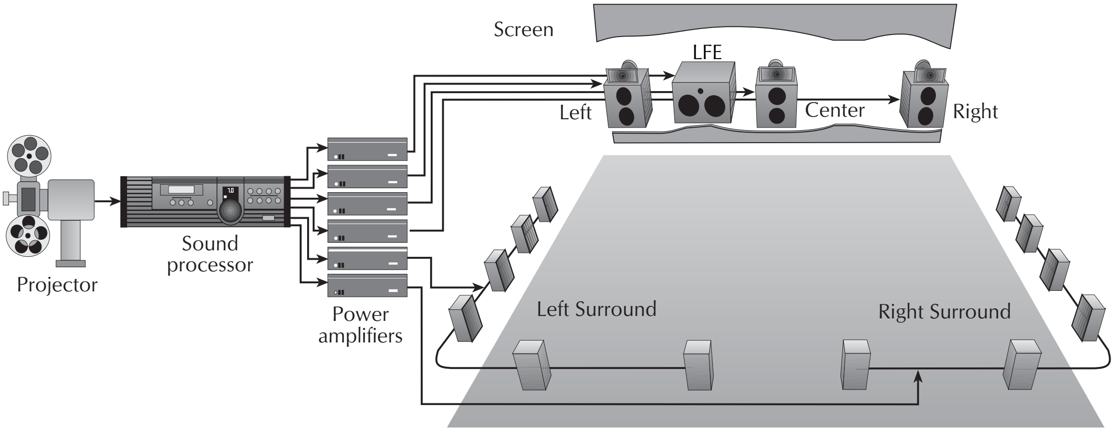

Screen
LFE
Left
Center
Right
Sound processor
Projector
Power amplifiers
Left Surround
Right Surround

Figure 45-4. The 5.1-channel format was first adopted for digital surround sound in the cinema.

conventional 5.1 playback systems, while cinema owners wishing to take advantage of the extra channel can equip their theaters with an additional decoder unit and power amplifiers. No additional loudspeakers are usually required, only rewiring the existing banks of surround loudspeakers, Fig. 45-5.

45.5 Surround Sound Comes Home

In 1982, recognizing the increasing popularity of watching VHS videotapes of theatrical movies in the home—and that the extra, matrix-encoded channels on Dolby Stereo movies were being transferred intact to their stereo (two-track) VHS versions—Dolby introduced the concept of surround sound in the home. It was dubbed Dolby Surround to differentiate it from the film process then still known as Dolby Stereo, and the term is still used today to identify any program material with stereo (two-track) soundtracks matrix-encoded for four-channel surround playback at the viewer's option.

Initially, Dolby developed a simple decoder circuit that passively derived just the surround channel from encoded VHS soundtracks, then licensed it to consumer electronics manufacturers. Later, in 1987, they introduced and began licensing a more sophisticated, true four-channel decoder, called Dolby Surround Pro Logic, with active steering and other features adapted from their professional cinema sound processors.

Pro Logic decoding, which began to be featured more and more in multichannel home playback products, heralded a new kind of home entertainment system, the home theater. During the late 1980s and well into the

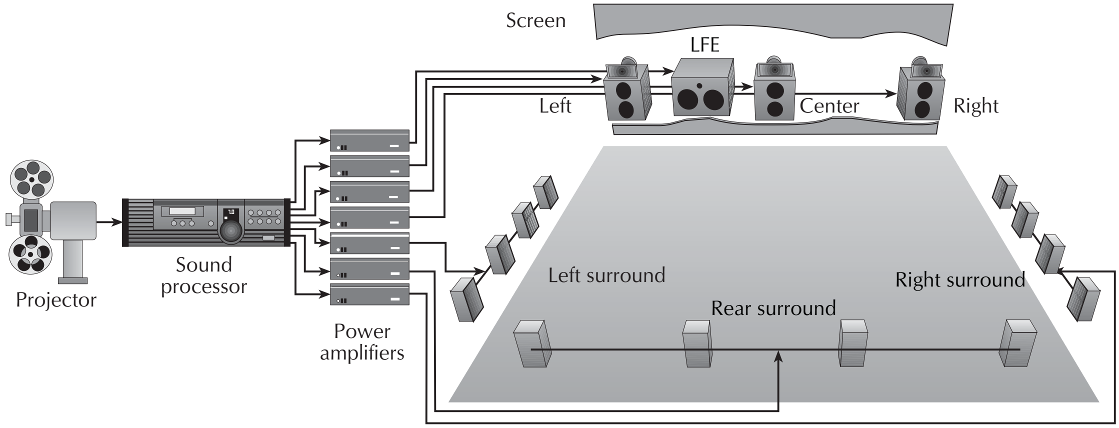

Screen
LFE
Left
Center
Right
Projector
Sound processor
Power amplifiers
Left surround
Right surround
Rear surround

Figure 45-5. Surround EX adds a third surround channel at the rear of the auditorium.

next decade, home theater was the fastest growing category of consumer electronics products, and today is enjoyed by many millions of consumers worldwide.

The appetite for Dolby Surround encoded programming increased accordingly, not only for videos of Dolby encoded movies, but also regular TV series, specials, and sports events. Content providers rose to the occasion by mixing more and more program material in Dolby Surround, confident that the material would play properly on any system, mono, stereo, or surround. And the Dolby Surround format soon extended to include the soundtracks of video and PC games, Fig. 45-6.

The purpose of a home theater system is to provide a convincing facsimile of what is heard in the cinema. In order to do that, the speakers should be placed as shown in Fig. 45-6, with three speakers across the front of the viewing area at about ear level, with the left and right speakers subtending a 45–60 degrees angle with the center seating position. A surround speaker goes to either side of the prime seating area well above ear level. Their relatively high placement helps to provide a diffuse surround soundfield, like that in a cinema, that does not call attention to itself.

When home theater was in its infancy, some pundits were skeptical that home listeners would put up with five loudspeakers in their living rooms (just as their ancestors had predicted failure for home stereo because two speakers were required). However, enough enthusiasts invested in the original, bulky home theater equipment to prompt loudspeaker manufacturers to develop sleek satellite/subwoofer systems that both eliminated most of the objections and lowered costs. While many devotees still assemble home theater systems from elaborate tower loudspeakers and other models, the majority of home listeners today opt for single-brand sub/sat systems.

Satellite/subwoofer systems take advantage of the fact that the lowest bass frequencies are nondirectional, that is, the ear cannot readily detect where bass sounds are coming from. As a result, these systems channel the low bass to a dedicated bass loudspeaker called a subwoofer. The subwoofer can usually be tucked out of the way, because its placement is not critical to reproducing the directionality of the original sound.

Because they are not required to reproduce low bass, the satellite loudspeakers can be compact, making them less intrusive and easier to place. Many systems use identical satellites for the left, center, right, and surround channels. This means that all loudspeakers have the same timbre, or tonal characteristic, which is desirable in a home theater system. Other systems provide identical satellites for left, center, and right, and somewhat different units (usually with respect to their radiating characteristic) for the surrounds. The surround loudspeakers should still be timbre-matched to the front loudspeakers.

45.6 Digital 5.1 in the Home

Much like Dolby's original analog film sound formats migrated into the home as Dolby Surround, Dolby Digital in the cinema provided a springboard for consumer formats with 5.1-channel digital surround. Beginning with laser discs in 1995, Dolby Digital 5.1 soon made its way to DVD, cable TV, DBS systems, digital TV broadcasting, and multimedia applications including video and PC games, Fig.45-7. DTS also entered the home market, although program material with DTS encoded soundtracks is found on relatively few DVD titles, and is unavailable via digital broadcast formats.

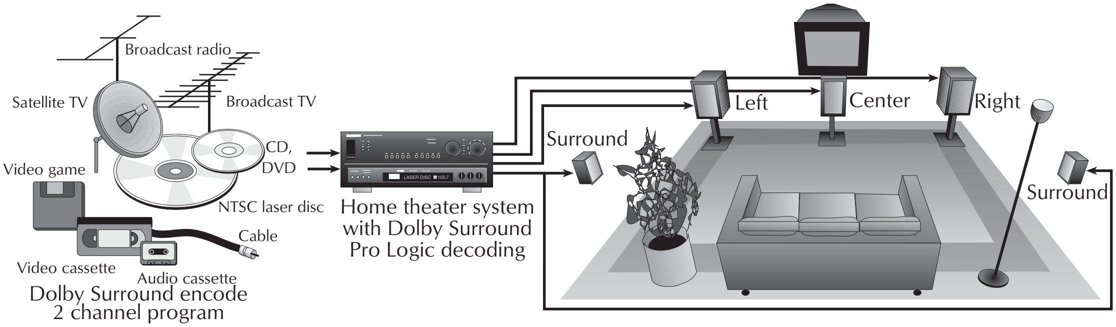


Figure 45-6. A four-channel home theater system equipped with Pro Logic decoding is configured much like a four-channel cinema system. Surprisingly, the need for at least five loudspeakers did not prove to be a deterrent, particularly as loudspeaker manufacturers developed compact and cost-effective satellite subwoofer systems

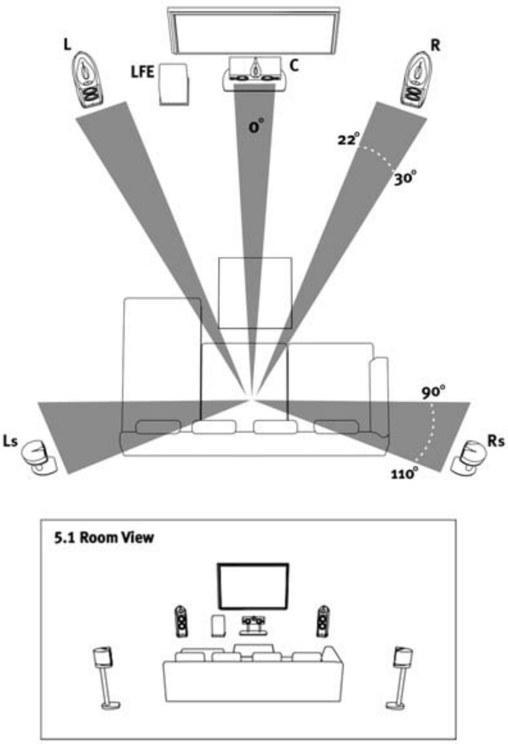


Figure 45-7. The speaker setup for digital 5.1-channel surround sound in the home is about the same as that for earlier four-channel surround. In a 5.1 setup, however, the left and right surround speakers are fed separate left and right surround information, where in a four-channel system they reproduce the same mono-surround information.

That film sound has been the starting point for 5.1 digital surround in the home has enabled the accumulation of invaluable experience in mixing, recording, and distributing multichannel digital audio. What's more, the widespread adoption of the Dolby Digital 5.1 format for consumer applications has resulted in the most direct link from program producer to home listener ever, giving the former unprecedented control over what the latter actually hears.

This is because the Dolby Digital bitstream carries not only the soundtrack as originally mixed, but also metadata, or data about the data, that can be used to control the home listener's Dolby Digital decoder. For example, while the same unrestricted multichannel audio content is delivered to every system, the consumer decoder can be instructed by metadata precisely how to downmix a 5.1-channel soundtrack for stereo or Pro Logic surround playback, or even mono playback. Metadata also allows the original mixers to pass onto the home decoder instructions that will, when the listener wishes, create a compressed version of the soundtrack on the fly for late-night viewing, when unrestricted dynamic range could bother family or neighbors.

45.7 More Home Theater Channels from Existing Content

The success of the 5.1 format, followed by the introduction of EX in 1999, led to the consumer electronics industry offering home theater listeners more channels by means of advanced matrix decoders and additional amplifiers in equipment such as audio/video receivers (AVRs). The first step was to offer decoding of the extra surround channel on DVDs of EX films for 6.1 playback, with the extra surround channel reproduced by a third surround speaker placed behind the listening area. A further variation soon followed, using two back surround speakers in a mono configuration reproducing the EX channel.

This 7.1 format, Fig. 45-8, as it is now known by the CE industry, has taken on greater popularity with the introduction of matrix technologies such as Dolby Pro Logic IIx and Harman-Kardon's Logic 7. These derive stereo back surround channels from regular 5.1, and even stereo, program sources, and today nearly all home theater AVRs in the $300 and up range are equipped for 7.1 playback.[*]

Deriving surround from stereo content became a practical and successful proposition with the introduction in 2000 of Dolby Pro Logic II, a 5.1 matrix-based decoding technology. Using a concept originally developed by audio pioneer Jim Fosgate, the decoder in essence seeks out surround cues occurring naturally in stereo content, such as ambience in music recordings. Pro Logic II can also be used to encode specific Left and right surround information up front onto stereo soundtracks, to achieve specific directional surround effects on playback with Pro Logic II decoding. This approach to delivering encoded surround content via stereo formats is replacing the older Dolby Surround technology for applications such as stereo broadcasting, and is used by some video game and console manufacturers as a practical and effective alternative to higher performance, but processing-power-hungry, digital 5.1 interactive audio.

---

[*] For home listeners not ready to commit to seven speakers in their living rooms, some 7.1 AVRs make it possible to use the extra two amplifier channels for stereo playback in another room, with the main system set up for standard 5.1 playback.

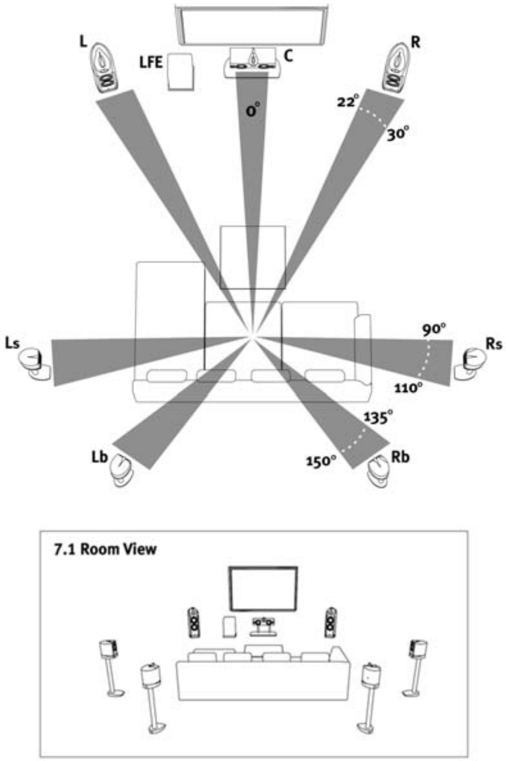


Figure 45-8. Speaker configuration for 7.1 home theater.

Dolby later expanded Pro Logic II technology to derive 7.1 surround (Pro Logic IIx, above), while competing systems from DTS, SRS, and others along with Dolby's have made surround sound from stereo content a standard feature in home theater systems and an increasing number of automobile sound systems. Many of these companies have also developed surround virtualizing technologies to provide a surround effect via just two stereo speakers. These technologies vary widely in their cost and effectiveness, but the best can provide quite startling results, albeit only in a limited listening "sweet spot," while the less sophisticated can provide a pleasant broadening of the stereo image over the stereo speakers built into TV sets.

45.8 The Question of Playback Level

When Dolby introduced surround on optical movie soundtracks in the late 1970s, it also introduced the concept of a reference playback level for both cinemas and the dubbing theaters where soundtracks are mixed. By calibrating both to the same reference level, the moviegoer hears in the cinema the same level as the director and sound designers heard when mixing the soundtrack. The objective—further supported by the introduction by Lucasfilm of its rigorous THX standards for playback quality—was for the moviegoer to experience the filmmakers' original intentions.

Over the past three decades, the standardization of playback level and other characteristics has rivaled surround itself as a significant improvement to the cinema experience. However, what about home playback of these same movie soundtracks via disc, tape, or broadcast?

Theoretically, to reproduce the cinema experience exactly, movies at home should be played at the same level as in the cinema. That level in the home, however, is both difficult to achieve and far too loud for most viewers. Most home movie viewers choose a level they find comfortable for dialogue intelligibility, which is substantially lower than cinema reference level. This results in a loss of impact compared to the cinema due to one of the peculiarities of human hearing: a loss of sensitivity to low and high frequencies that increases as playback level decreases. Indeed at lower playback levels, low-level detail, such as ambience in the surround channels, can disappear altogether.

So-called loudness controls that attempt to compensate for this effect have been incorporated in home playback equipment for many years. They boost low and high frequencies, usually based in some way on the famous equal loudness curves originally published by Fletcher and Munson in 1933 and updated over the years, Fig. 45-9. However, until recently, there has been no practical way to relate the actions of these controls to the actual playback level the listener has chosen. For the most part, therefore, they have been largely ineffective.

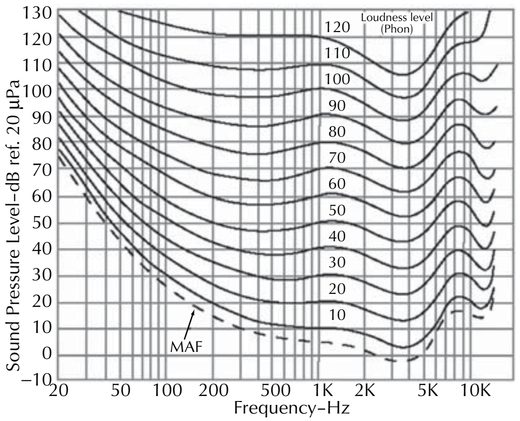


Figure 45-9. The equal-loudness contours. Illustration courtesy Syn-Aud-Con.

What has been needed is a way to establish a playback reference level based on the actual measured acoustic output of the listener's playback system. The scientifically correct compensation could then be applied based on the listener's preferred listening level relative to the reference level. The lower the level, the more compensation would be applied.

Measuring the acoustic output of a home system was once beyond the realm of practicality. Today, however, it happens all the time in home theater systems featuring audio-video receivers (AVR) equipped with automatic loudspeaker balancing. The majority of mid- to high-end AVRs has built-in noise generators and provides a microphone making it possible to measure the loudness of each loudspeaker from the listening position and automatically adjust it to match the others. Taking advantage of this built-in feature, several technologies, including Dolby Volume, THX Spectral Balancing, and Audyssey Dynamic EQ, have been recently introduced that make it possible to establish a reference playback level in the home, and apply appropriate loudness compensation at lower levels. It is now possible to automatically achieve at any level the same balance of low, middle, and high frequencies, and of main to surround channels, as at reference level, bringing home reproduction that much closer to what sound mixers achieve in the dubbing theater or music studio.

45.9 What's Next for Surround Sound?

5.1 remains the standard for film-based cinema, with EX used with some regularity for big epic and sci-fi films. This is in part because of cost and complexity issues, in part because the movie industry is pouring its resources into converting to digital cinema, and in part because of the industry's overall industry satisfaction with 5.1 both artistically and with respect to cost-effectiveness. After all, as stereo pioneer Harvey Fletcher put it way back in the 1940s,

> Stereophonic systems do not consist of two, three, or any other fixed number of channels. There [only] must be sufficient of these to give a good illusion of an infinite number.

On the other hand, digital cinema content is *capable* of delivering twenty or more channels of uncompressed PCM audio. Even with the increasing number of digital cinema installations, however, 5.1 remains the standard for digital releases. Not only is equipping cinemas with more playback channels expensive, but so is mixing soundtracks with more channels, particularly since mixing is one of the final postproduction steps when time may be running out. Moreover, whether to use the additional channels—and if so, which ones—is wide open. As the SMPTE puts it in its standard for digital cinema channel mapping (428-3-2006), "This standard is not intended to define the suitability of these channels to a particular track, nor to specify that all the channels described herein will be used," Table 45-1.

Because the movie industry has not yet ventured into the realm of more than 5.1 discrete channels, it continues to deliver 5.1 content for broadcast and video disc release. But that isn't stopping the consumer electronics industry from experimenting with 7.1 discrete audio content, just pioneered matrix 7.1 playback. Both Blu-ray Disc and HD DVD have ample storage capability for more channels, whether uncompressed PCM, or using lossless or lossy coding technologies offered by both Dolby and DTS. A few hardy pioneers are going back to movie soundtrack stems and remixing titles in discrete 7.1 for high-definition disc release, with the blessing and supervision of the film's producers.

Playback in 7.1 discrete depends on home theater AVRs equipped with appropriate decoders, and both players and AVRs equipped with HDMI 1.3 connectivity, which is far from universal. It's impossible to predict just what will happen, but it's conceivable that the consumer electronics industry, whose surround sound technology so far has mostly migrated from the movie industry, may wind up ahead in the multichannel race—assuming consumers buy into it, of course.

Perhaps more significant than the potential for more channels, however, are the rapidly expanding opportunities for delivering 5.1 content. New broadcast standards developed to take advantage of new, more efficient video and audio codecs alike all feature 5.1 capability. These new codecs are fostering new delivery methods, such as IPTV, for 5.1 audio. And the future is likely to offer 5.1 download opportunities, such as Apple's pioneering effort enabling the purchase and rentals of downloaded high-definition movies with Dolby Digital 5.1 audio. Indeed, there are those predicting that in the foreseeable future the Internet will overtake discs as the prime conduit for movies and other surround content into the home.

Regardless of what the future brings, however, surround sound has come a very long way already, from rarefied and costly magnetic sound in cinemas 50 years ago, to home theater audio systems costing as little as a few hundred dollars today. Involving the viewer is what surround is all about, and there's no doubt that moviegoers and home viewers alike prefer—even demand—the surround experience in ever-increasing numbers, for every possible entertainment medium.

Table 45-1. Channel Definitions for Digital Cinema (SMPTE Standard 428-3-2006). So Far, However, 5.1 Remains the Norm for Digital as Well as Film-Based

Left. A loudspeaker position behind the screen to the far left edge, horizontally, of the screen center as viewed from the seating area.

Center. A loudspeaker position behind the screen corresponding to the horizontal center of the screen as viewed from the seating area.

Right. A loudspeaker position behind the screen to the far right edge, horizontally, of the screen center as viewed from the seating area.

LFE screen. A band-limited low-frequency-only loudspeaker position at the screen end of the room. Also referred to as the subwoofer channel.

Left surround. An array of loudspeakers positioned along the left side of the room starting approximately ⅓ of the distance from the screen to the back wall.

Right surround. An array of loudspeakers positioned along the right side of the room starting approximately ⅓ of the distance from the screen to the back wall.

Center surround. A loudspeaker(s) position on the back wall of the room centered horizontally.

Left center. A loudspeaker position midway between the center of the screen and the left edge of the screen.

Right center. A loudspeaker position midway between the center of the screen and the right edge of the screen.

LFE 2. A band-limited low-frequency-only loudspeaker.

Vertical height front. A loudspeaker(s) position at the vertical top of the screen. A single channel would be at the center of the screen horizontally. Dual channels may be positioned at the vertical top of the screen and in the left center and right center horizontal positions. Tri-channel may be positioned at the vertical top of the screen in the left, center and right horizontal positions.

Top center surround. A loudspeaker position in the center of the seating area in both the horizontal and vertical planes directly above the seating area.

Left wide. A loudspeaker position outside the screen area far left front in the room.

Right wide. A loudspeaker position outside the screen area far right front in the room.

Rear surround left. A loudspeaker position on the back wall of the room to the left horizontally.

Rear surround right. A loudspeaker position on the back wall of the room to the right horizontally.

Left surround direct. A loudspeaker position on the left wall for localization as opposed to the diffuse array.

Right surround direct. A loudspeaker position on the right wall for localization as opposed to the diffuse array.

Hearing impaired. A dedicated audio channel optimizing dialoue intelligibility for the hearing impaired.

Narration. A dedicated narration channel describing the films' events for the visually impaired.

This page intentionally left blank

Part 7

Measurements

This page intentionally left blank

Chapter 46

Test and Measurement

by Pat Brown

46.1 Test and Measurement 1607
46.1.1 Why Test? 1607
46.2 Electrical Testing 1607
46.3 Acoustical Testing 1607
46.3.1 Sound Level Measurements 1608
46.3.1.1 Conclusion 1610
46.3.2 Detailed Sound Measurements 1611
46.3.2.1 Sound Persistence in Enclosed Spaces 1611
46.3.2.2 Amplitude versus Time 1611
46.3.2.3 The Transfer Function 1613
46.3.3 Measurement Systems 1613
46.3.3.1 Alternate Perspectives 1613
46.3.4 Testing Methods 1614
46.3.4.1 FFT Measurements 1615
46.3.4.2 Dual-Channel FFT 1616
46.3.4.3 Maximum-Length Sequence 1616
46.3.4.4 Time-Delay Spectrometry (TDS) 1616
46.3.5 Preparation 1617
46.3.5.1 Dissecting the Impulse Response 1618
46.3.5.2 The Envelope-Time Curve 1618
46.3.5.3 A Global Look 1618
46.3.5.4 Time Window Length 1618
46.3.5.5 Acoustic Wavelengths 1619
46.3.5.6 Microphone Placement 1619
46.3.5.7 Estimate the Critical Distance DC 1620
46.3.5.8 Common Factors to All Measurement Systems 1620
46.3.5.9 Data Windows 1622
46.3.5.10 A Methodical Approach 1623
46.3.6 Human Perception 1625
46.3.6.1 Percentage Articulation Loss of Consonants- (%Alcons) 1625
46.3.6.2 Speech Transmission Index—(STI) 1627
46.3.7 Polarity 1627
46.3.7.1 The Battery Test 1627
46.3.7.2 Polarity Testers 1627
46.3.7.3 Impulse Response Tests 1628
46.4 Conclusion 1628
References 1629
Bibliography 1629

This page intentionally left blank

46.1 Test and Measurement

Technological advancements in the last two decades have given us a variety of useful measurement tools, and most manufacturers of these instruments provide specialized training on their use. This chapter will examine some principles of test and measurement that are common to virtually all measurement systems. If the measurer understands the principles of measurement, then most any of the mainstream measurement tools will suffice for the collection and evaluation of data. The most important prerequisite to performing meaningful sound system measurements is that the measurer has a solid understanding of the basics of audio and acoustics. The question "How do I perform a measurement?" can be answered much more easily than "What should I measure?" This chapter will touch on both, but readers will find their measurements skills will relate directly to their understanding of the basic physics of sound and the factors that produce good sound quality. The whole of this book will provide much of the required information.

46.1.1 Why Test?

Sound systems must be tested to assure that all components are functioning properly. The test and measurement process can be subdivided into two major categories—electrical tests and acoustical tests. Electrical testing mainly involves voltage and impedance measurements made at component interfaces. Current can also be measured, but since the setup is inherently more complex it is usually calculated from knowledge of the voltage and impedance using Ohm's Law. Acoustical tests are more complex by nature, but share the same fundamentals as electrical tests in that some time varying quantity (usually pressure) is being measured. The main difference between electrical and acoustical testing is that the interpretation of the latter must deal with the complexities of 3D space, not just amplitude versus time at one point in a circuit. In this chapter we will define a loudspeaker system as a number of components intentionally combined to produce a system that may then be referred to as a loudspeaker. For example, a woofer, dome tweeter, and crossover network are individual components, but can be combined to form a loudspeaker system. Testing usually involves the measurement of systems, although a system might have to be dissected to fully characterize the response of each component.

46.2 Electrical Testing

There are numerous electrical tests that can be performed on sound system components in the laboratory. The measurement system must have specifications that exceed the equipment being measured. Field testing need not be as comprehensive and the tests can be performed with less sophisticated instrumentation. The purpose for electrical field testing includes:

1. To determine if all system components are operating properly.
2. To diagnose electrical problems in the system, which are usually manifested by some form of distortion.
3. To establish a proper gain structure.

Electrical measurements can aid greatly in establishing the proper gain structure of the sound system. Electrical test instruments that the author feels are essential to the audio technician include:

- ac voltmeter.
- ac millivoltmeter.
- Oscilloscope.
- Impedance meter.
- Signal generator.
- Polarity test set.

It is important to note that most audio products have on-board metering and/or indicators that may suffice for setting levels, making measurements with stand-alone meters unnecessary. Voltmeters and impedance meters are often only necessary for troubleshooting a nonworking system, or checking the accuracy and calibration of the on-board metering.

There are a number of currently available instruments designed specifically for audio professionals that perform all of the functions listed. These instruments need to have bandwidths that cover the audible spectrum. Many general purpose meters are designed primarily for ac power circuits and do not fit the wide bandwidth requirement.

More information on electrical testing is included in the chapter on gain structure. The remainder of this chapter will be devoted to the acoustical tests that are required to characterize loudspeakers and rooms.

46.3 Acoustical Testing

The bulk of acoustical measurement and analysis today is being performed by instrumentation that includes or is controlled by a personal computer. Many excellent systems are available, and the would-be measurer should select the one that best fits their specific needs. As with loudspeakers, there is no clear-cut best choice or one size fits all instrument. Fortunately an understanding of the principles of operating one analyzer can usually be applied to another after a short indoctrination period. Measurement systems are like rental cars—you know what features are there; you just need to find them. In this chapter I will attempt to provide a sufficient overview of the various approaches to allow the reader to investigate and select a tool to meet his or her measurement needs and budget. The acoustical field testing of sound reinforcement systems mainly involves measurements of the sound pressure fluctuations produced by a loudspeaker(s) at various locations in the space. Microphone positions are selected based on the information that is needed. This could be the on-axis position of a loudspeaker for system alignment purposes, or a listener seat for measuring the clarity or intelligibility of the system. Measurements must be made to properly calibrate the system, which can include loudspeaker crossover settings, equalization, and the setting of signal delays. Acoustic waveforms are complex by nature, making them difficult to describe with one number readings for anything other than broadband level.

46.3.1 Sound Level Measurements

Sound level measurements are fundamental to all types of audio work. Unfortunately, the question “How loud is it?” does not have a simple answer. Instruments can easily measure sound pressures, but there are many ways to describe the results in ways relevant to human perception. Sound pressures are usually measured at a discrete listener position. The sound pressure level may be displayed as is, integrated over a time interval, or frequency weighted by an appropriate filter. Fast meter response times produce information about peaks and transients in the program material, while slow response times yield data that correlates better with the perceived loudness and energy content of the sound.

A sound level meter consists of a pressure sensitive microphone, meter movement (or digital display), and some supporting circuitry, Fig. 46-1. It is used to observe the sound pressure on a moment-by-moment basis, with the pressure displayed as a level in decibels. Few sounds will measure the same from one instant to the next. Complex sounds such as speech and music will vary dramatically, making their level difficult to describe without a graph of level versus time, Fig. 46-2. A sound level meter is basically a voltmeter that operates in the acoustic domain.

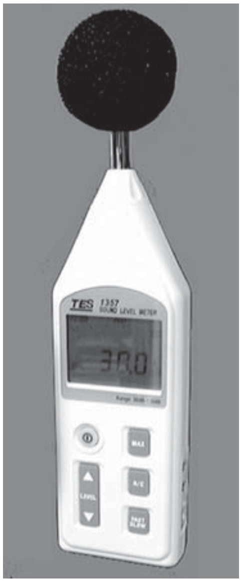


Figure 46-1. A sound level meter is basically a voltmeter that operates in the acoustic domain. Courtesy Galaxy Audio.

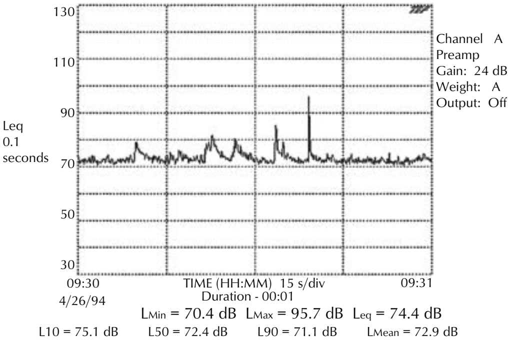


Figure 46-2. A plot of sound level versus time is the most complete way to record the level of an event. Courtesy Gold Line.

Sound pressure measurements are converted into decibels ref. 0.00002 pascals. See Chapter 2, Fundamentals of Audio and Acoustics, for information about the decibel. Twenty micropascals are used as the reference because it is the threshold of pressure sensitivity for humans at midrange frequencies. Such measurements are referred to as *sound pressure level* or L_P (level of sound pressure) measurements, with L_P gaining acceptance among audio professionals because it is easily distinguished from L_W (sound power level) and L_I (sound intensity level) and a number of other L_X metrics used to describe sound levels. Sound pressure level is measured at a single point (the microphone position). Sound power measurements must consider all of the radiated sound from a device and sound intensity measurements must consider the sound power flowing through an area. Sound power and sound intensity measurements are usually performed by acoustical laboratories rather than in the field, so neither is considered in this chapter. All measurements described in this chapter will be measurements of sound pressures expressed as levels in dB ref. 0.00002 Pa.

Sound level measurements must usually be processed for the data to correlate with human perception. Humans do not hear all frequencies with equal sensitivity, and to complicate things further our response is dependent on the level that we are hearing. The well-known Fletcher-Munson curves describe the frequency/level characteristics for an average listener, see Chapter 2. Sound level measurements are passed through weighting filters that make the meter "hear" with a response similar to a human. Each scale correlates with human hearing sensitivity at a different range of levels. For a sound level measurement to be meaningful, the weighting scale that was used must be indicated, in addition to the response time of the meter. Here are some examples of meaningful (if not universally accepted) expressions of sound level:

- The system produced an $L_P = 100$ dBA (slow response) at mix position.
- The peak sound level was $L_A = 115$ dB at my seat.
- The average sound pressure level was 100 dBC at 30 ft.
- The loudspeaker produced a continuous L_P of 100 dB at one meter (assumes no weighting used).
- The equivalent sound level L_{EQ} was 90 dBA at the farthest seat.

Level specifications should be stated clearly enough to allow someone to repeat the test from the description given. Because of the large differences between the weighting scales, it is meaningless to specify a sound level without indicating the scale that was used. An event that produces an $L_P = 115$ dB using a C scale may only measure as an $L_P = 95$ dB using the A scale.

The measurement distance should also be specified (but rarely is). Probably all sound reinforcement systems produce an $L_P = 100$ dB at some distance, but not all do so at the back row of the audience!

L_{pk} is the level of the highest instantaneous peak in the measured time interval. Peaks are of interest because our sound system components must be able to pass them without clipping them. A peak that is clipped produces high levels of harmonic distortion that degrade sound quality. Also, clipping reduces the crest factor of the waveform, causing more heat to be generated in the loudspeaker causing premature failure. Humans are not extremely sensitive to the loudness of peaks because our auditory system integrates energy over time with regard to loudness. We are, unfortunately, susceptible to damage from peaks, so they should not be ignored. Research suggests that it takes the brain about 35 ms to process sound information (frequency-dependent), which means that sound events closer together than this are blended together with regard to loudness. This is why your voice sounds louder in a small, hard room. It is also why the loudness of the vacuum cleaner varies from room to room. Short interval reflections are integrated with the direct sound by the ear/brain system. Most sound level meters have slow and fast settings that change the response time of the meter. The slow setting of most meters indicates the approximate root-mean-square sound level. This is the effective level of the signal, and should correlate well with its perceived loudness.

A survey of audio practitioners on the Syn-Aud-Con e-mail discussion group revealed that most accept an $L_P = 95$ dBA (slow response) as the maximum acceptable sound level of a performance at any listener seat for a broad age group audience. The A weighting is used because it considers the sound level in the portion of the spectrum where humans are most easily annoyed and damaged. The slow response time allows the measurement to ignore short duration peaks in the program. A measurement of this type will not indicate true levels for low-frequency information, but it is normally the mid-frequency levels that are of interest.

There exist a number of ways to quantify sound levels that are measured over time. They include:

- $\boldsymbol{L_{PK}}$—the maximum instantaneous peak recorded during the span.
- $\boldsymbol{L_{EQ}}$—the equivalent level (the integrated energy over a specified time interval).
- $\boldsymbol{L_N}$—where L is the level exceeded N percent of the time.
- $\boldsymbol{L_{DEN}}$—a special scale that weights the gathered sound levels based on the time of day. DEN stands for day-evening-night.
- ***DOSE—a measure of the total sound exposure.***

A variety of instruments are available to measure sound pressure levels, ranging from the simple sound level meter (SLM) to sophisticated data-logging equipment. SLMs are useful for making quick checks of sound levels. Most have at least an A- and C-weighting scale, and some have octave band filters that allow band-limited measurements. A useful feature on an SLM is an output jack that allows access to the measured data in the form of an ac voltage. Software applications are available that can log the meter's response versus time and display the results in various ways. A plot of sound level versus time is the most complete way to record the level of an event. Fig. 46-2 is such a measurement. Note that a start time and stop time are specified. Such measurements usually provide statistical summaries for the recorded data. An increasing number of venues monitor the levels of performing acts in this manner due to growing concerns over litigation about hearing damage to patrons. SLMs vary dramatically in price, depending on quality and accuracy.

All sound level meters provide accurate indications for relative levels. For absolute level measurements a calibrator must be used to calibrate the measurement system. Many PC-based measurement systems have routines that automate the calibration process. The calibrator is placed on the microphone, Fig. 46-3, and the calibrator level (usually 94 or 114 dB ref. 20 µPa) is entered into a data field. The measurement tool now has a true level to use as a reference for displaying measured data.

Noise criteria ratings provide a one-number specification for allowable levels of ambient noise. Sound level measurements are performed in octave bands, and the results are plotted on the chart shown in Fig. 46-4. The NC rating is read on the right vertical axis. Note that the NC curve is frequency-weighted. It permits an increased level of low-frequency noise, but becomes more stringent at higher frequencies. A sound system specification should include an NC rating for the space, since excessive ambient noise will reduce system clarity and require additional acoustic gain. This must be considered when designing the sound system. Instrumentation is available to automate noise criteria measurements.

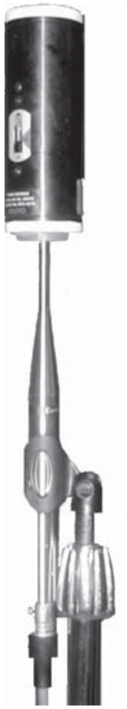


Figure 46-3. A calibrator must be fitted with a disc to provide a snug fit to the microphone. Most microphone manufacturers can provide the disc.

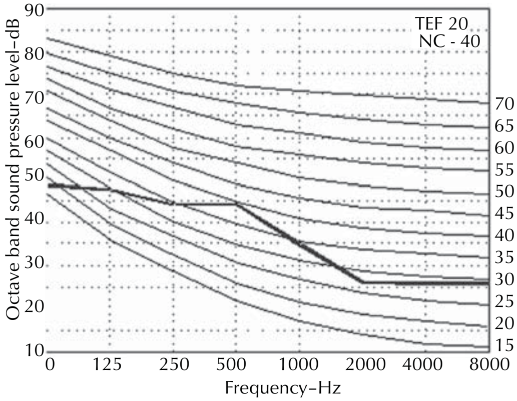


Figure 46-4. A noise criteria specification should accompany a sound system specification.

46.3.1.1 Conclusion

Stated sound level measurements are often so ambiguous as to become meaningless. When stating a sound level, it is important to indicate:

1. The sound pressure level.
2. Any weighting scale used.
3. Meter response time (fast, slow or other).
4. The distance or location at which the measurement was made.
5. The type of program measured (i.e., music, speech, ambient noise).

Some correct examples:

- "The house system produced 90 dBA-Slow in section C for broadband program."

- "The monitor system produced 105 dBA-Slow at the performer's head position for broadband program."
- "The ambient noise with room empty was NC-35 with HVAC running."

In short, if you read the number and have to request clarification then sufficient information has not been given. As you can see, one-number SPL ratings are rarely useful.

All sound technicians should own a sound level meter, and many can justify investment in more elaborate systems that provide statistics on the measured sound levels. From a practical perspective, it is a worthwhile endeavor to train one's self to recognize various sound levels without a meter, if for no other reason than to find an exit in a venue where excessive levels exist.

46.3.2 Detailed Sound Measurements

The response of a loudspeaker or room must be measured with appropriate frequency resolution to be characterized. It is also important for the measurer to understand what the appropriate response should be. If the same criteria were applied to a loudspeaker as to an electronic component such as a mixer, the optimum response would be a flat (minimal variation) magnitude and phase response at all frequencies within the required pass band of the system. In reality, we are usually testing loudspeakers to make sure that they are operating at their fullest potential. While flat magnitude and phase response are a noble objective, the physical reality is that we must often settle for far less in terms of accuracy. Notwithstanding, even with their inherent inaccuracies, many loudspeakers do an outstanding job of delivering speech or music to the audience. Part of the role of the measurer is to determine if the response of the loudspeaker or room is inhibiting the required system performance.

46.3.2.1 Sound Persistence in Enclosed Spaces

Sound system performance is greatly affected by the sound energy persistence in the listening space. One metric that is useful for describing this characteristic is the reverberation time, T_{30}. The T_{30} is the time required for an interrupted steady-state sound source to decay to inaudibility. This will be about 60 dB of decay in most auditoriums with controlled ambient noise floors. The T_{30} designation comes from the practice of measuring 30 dB of decay and then doubling the time interval to get the time required for 60 dB of decay. A number of methods exist for determining the T_{30}, ranging from simple listening tests to sophisticated analytical methods. Fig. 46-5 shows a simple gated-noise test that can provide sufficient accuracy for designing systems. The bursts for this test can be generated with a WAV editor. Bursts of up to 5 seconds for each of eight octave bands should be generated. Octave band-limited noise is played into the space through a low directivity loudspeaker. The noise is gated on for one second and off for 1 second. The room decay is evaluated during the off span. If it decays completely before the next burst, the T_{30} is less than one second. If not, the next burst should be on for 2 seconds and off for 2 seconds. The measurer simply keeps advancing to the next track until the room completely decays in the off span, Figs. 46-6, 46-7, and 46-8. The advantages of this method include:

1. No sophisticated instrumentation is required.
2. The measurer is free to wander the space.
3. The nature of the decaying field can be judged.
4. A group can perform the measurement.

A test of this type is useful as a prelude to more sophisticated techniques.

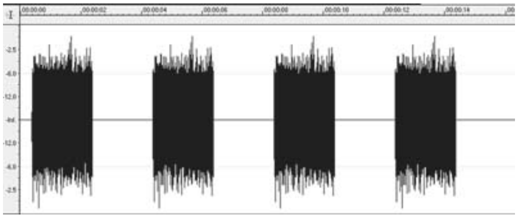


Figure 46-5. Level versus time plot of a one-octave band gated burst (2-second duration).

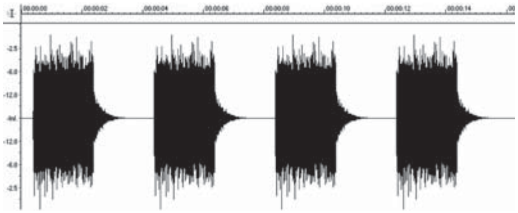


Figure 46-6. A room with RT_{60} <2 seconds.

46.3.2.2 Amplitude versus Time

Fig. 46-9 shows an audio waveform displayed as amplitude versus time. This representation is especially meaningful to humans since it can represent the motion of the eardrum about its resting position. The waveform shown is of a male talker recorded in an anechoic (echo-free)

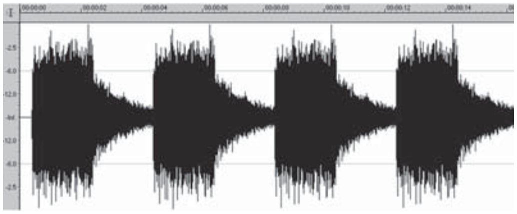


Figure 46-7. A room with RT_{60} >2 seconds.

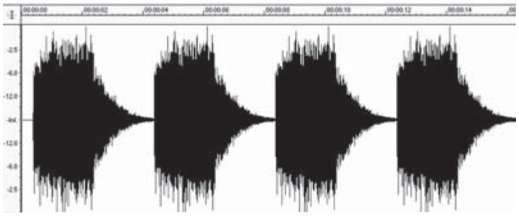


Figure 46-8. A room with RT_{60} = 2 seconds.

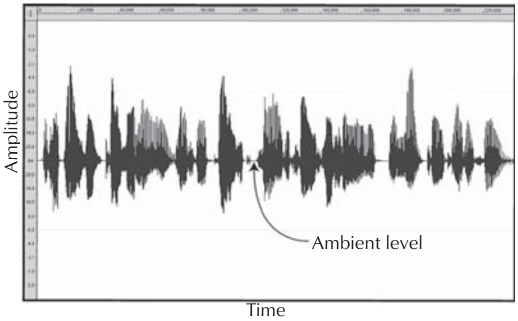


Figure 46-9. Amplitude versus time plot of a male talker made in an anechoic environment.

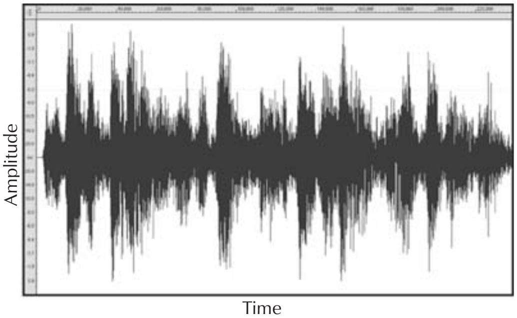


Figure 46-10. The voice waveform after encoding with the room response.

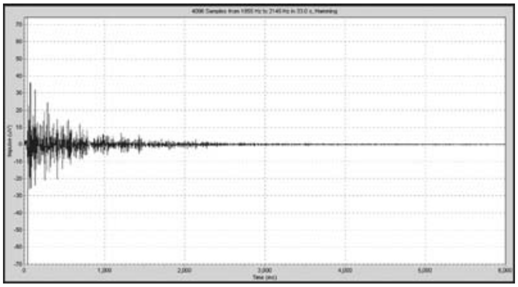


Figure 46-11. The impulse response of the acoustic environment.

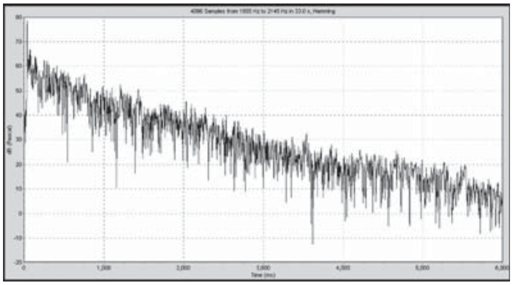


Figure 46-12. The envelope-time curve (ETC) of the same environment. It can be derived from the impulse response.

environment. The 0 line represents the ambient (no signal) state of the medium being modulated. This would be ambient atmospheric pressure for an acoustical wave, or zero volts or a dc offset for an electrical waveform measured at the output of a system component.

Fig. 46-10 shows the same waveform, but this time played over a loudspeaker into a room and recorded. The waveform has now been encoded (convolved) with the response of the loudspeaker and room. It will sound completely different than the anechoic version.

Fig. 46-11 shows an impulse response and Fig. 46-12 shows the envelope-time curve (ETC) of the loudspeaker and room. It is essentially the difference between Fig. 46-9 and Fig. 46-10 that fully characterizes any effect that the loudspeaker or room has on the electrical signal fed to the loudspeaker and measured at that point in space. Most measurement systems attempt to measure the impulse response, since knowledge of the impulse response of a system allows its effect on any signal passing through it to be determined, assuming the system is linear and time invariant. This effect is called the transfer function of the system and includes both magnitude (level) and phase (timing) information for each frequency in the pass band. Both the loudspeaker and room can be considered filters that the energy must pass through en route to the listener. Treating them as filters allows their responses to be measured and displayed, and provides an objective benchmark to evaluate their effect. It also opens loudspeakers and rooms to evaluation by electrical network analysis methods, which are generally more widely known and better developed than acoustical measurement methods.

46.3.2.3 The Transfer Function

The effect that a filter has on a waveform is called its *transfer function*. A transfer function can be found by comparing the input signal and output signal of the filter. It matters little if the filter is an electronic component, loudspeaker, room, or listener. The time domain behavior of a system (impulse response) can be displayed in the frequency domain as a spectrum and phase (transfer function). Either the time or frequency description fully describes the filter. Knowledge of one allows the determination of the other. The mathematical map between the two representations is called a transform. Transforms can be performed at amazingly fast speeds by computers. Fig. 46-13 shows a domain chart that provides a map between various representations of a system's response. The measurer must remember that the responses being measured and displayed on the analyzer are dependent on the test stimulus used to acquire the response. Appropriate stimuli must have adequate energy content over the pass band of the system being measured. In other words, we can't measure a subwoofer using a flute solo as a stimulus. With that criteria met, the response measured and displayed on the analyzer is independent of the program material that passes through a linear system. Pink noise and sine sweeps are common stimuli due to their broadband spectral content. In other words, the response of the system doesn't change relative to the nature of the program material. For a linear system, the transfer function is a summary that says, "If you put energy into this system, this is what will happen to it."

The domain chart provides a map between various methods of displaying the system's response. The utility of this is that it allows measurement in either the time or frequency domain. The alternate view can be determined mathematically by use of a transform. This allows frequency information to be determined with a time domain measurement, and time information to be determined by a frequency domain measurement. This important inverse relationship between time and frequency can be exploited to yield many possible ways of measuring a system and/or displaying its response. For instance, a noise immunity characteristic not attainable in the time domain may be attainable in the frequency domain. This information can then be viewed in the time domain by use of a transform. The Fourier Transform and its inverse are commonly employed for this purpose. Measurement programs like Arta can display the signal in either domain, Fig. 46-14.

46.3.3 Measurement Systems

Any useful measurement system must be able to extract the system response in the presence of noise. In some applications, the signal-to-noise requirements might actually determine the type of analysis that will be used. Some of the simplest and most convenient tests have poor signal-to-noise performance, while some of the most complex and computationally demanding methods can measure under almost any conditions. The measurer must choose the type of analysis with these factors in mind. It is possible to acquire the impulse response of a filter without using an impulse. This is accomplished by feeding a known broadband stimulus into the filter and reacquiring it at the output. A complex comparison of the two signals (mathematical division) yields the transfer function, which is displayed in the frequency domain as a magnitude and phase or inverse-transformed for display in the time domain as an impulse response. The impulse response of a system answers the question, "If I feed a perfect impulse into this system, when will the energy exit the system?" A knowledge of "when" can characterize a system. After transformation, the spectrum or frequency response is displayed on a decibel scale. A phase plot shows the phase response of the device-under-test, and any phase shift versus frequency becomes apparent. If an impulse response is a measure of when, we might describe a frequency response as a measure of what. In other words, "If I input a broadband stimulus (all frequencies) into the system, what frequencies will be present at the output of the system and what will their phase relationship be?" A transfer function includes both magnitude and phase information.

46.3.3.1 Alternate Perspectives

The time and frequency views of a system's response are mutually exclusive. By definition the time period of a periodic event is

$$T = \frac{1}{f} \tag{46-1}$$

where,
T is time in seconds,
f is frequency in hertz.

Since time and frequency are reciprocals, a view of one excludes a view of the other. Frequency information cannot be observed on an impulse response plot, and

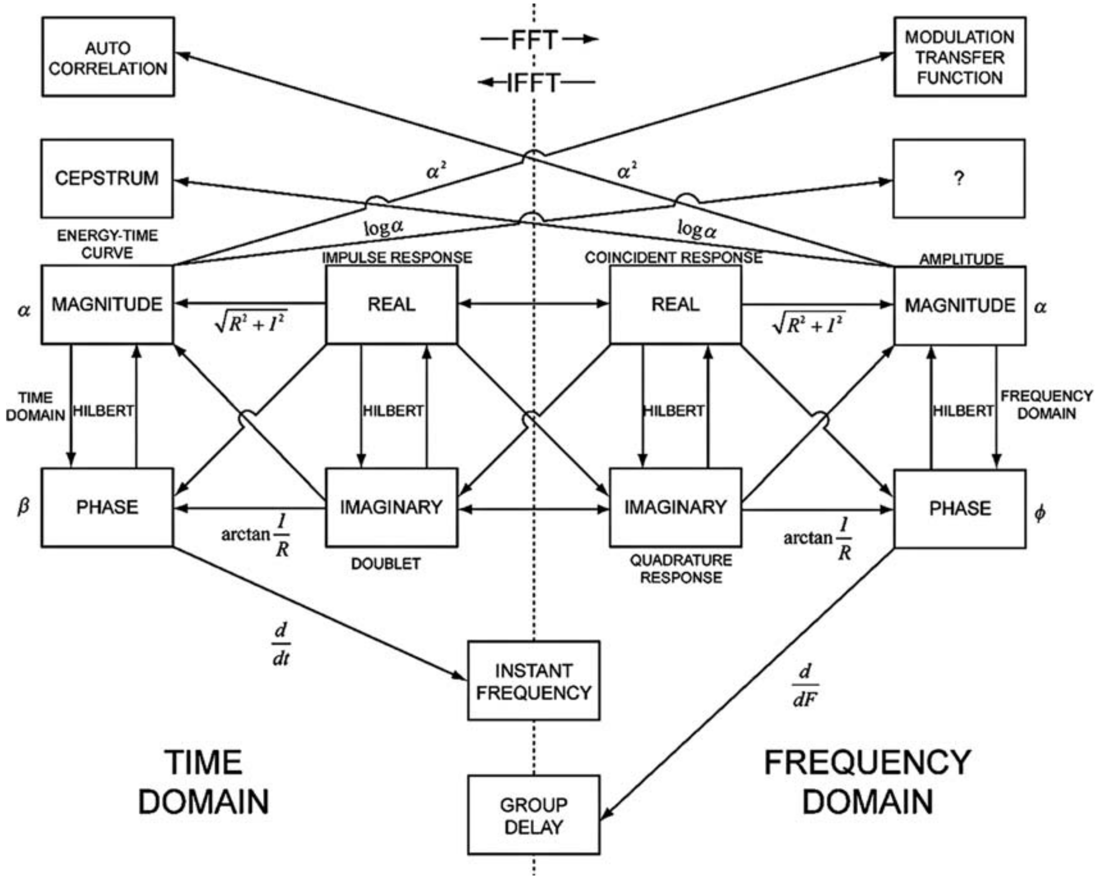

FFT
IFFT
AUTO CORRELATION
MODULATION TRANSFER FUNCTION
CEPSTRUM
?
α^2
α^2
$\log\alpha$
$\log\alpha$
ENERGY-TIME CURVE
IMPULSE RESPONSE
COINCIDENT RESPONSE
AMPLITUDE
α
MAGNITUDE
REAL
REAL
MAGNITUDE
α
$\sqrt{R^2+I^2}$
$\sqrt{R^2+I^2}$
TIME DOMAIN
HILBERT
HILBERT
HILBERT
HILBERT
FREQUENCY DOMAIN
β
PHASE
IMAGINARY
IMAGINARY
PHASE
ϕ
$\arctan\frac{I}{R}$
$\arctan\frac{I}{R}$
DOUBLET
QUADRATURE RESPONSE
$\frac{d}{dt}$
INSTANT FREQUENCY
$\frac{d}{dF}$
TIME DOMAIN
FREQUENCY DOMAIN
GROUP DELAY

Figure 46-13. The domain chart provides a map between various representations of a system response. Courtesy Brüel and Kjaer.

time information can't be observed on a magnitude/phase plot. Any attempt to view both simultaneously will obscure some of the detail of both. Modern analyzers allow the measurer to switch between the time and frequency perspectives to extract information from the data.

46.3.4 Testing Methods

Compared to other components in the sound system, the basic design of loudspeakers and compression drivers has changed relatively little in the last 50 years. At over a half-century since their invention, we are still pushing air with pistons driven by voice coils suspended in magnetic fields. But the methods for measuring their performance have improved steadily since computers can now efficiently perform digital sampling and signal processing, and execute transforms in fractions of a second. Extremely capable measurement systems are now accessible and affordable to even the smallest manufacturers and individual audio practitioners. A common attribute of systems suitable for loudspeaker testing is the ability to make reflection-free measurements indoors, without the need for an anechoic chamber. Anechoic measurements in live spaces can be accomplished by the use of a time window that allows the analyzer to collect the direct field response of the loudspeaker while ignoring room reflections. Conceptually, a time window can be thought of as an accurate switch that can be closed as the desired waves pass the microphone and opened prior to the arrival of undesirable reflections from the environment. A number of implementations exist, each with its own set of advantages and drawbacks. The potential buyer must understand the trade-offs and choose a system that offers the best set of compromises for the intended application. Parameters of interest include signal-to-noise ratios, speed, resolution, and price.

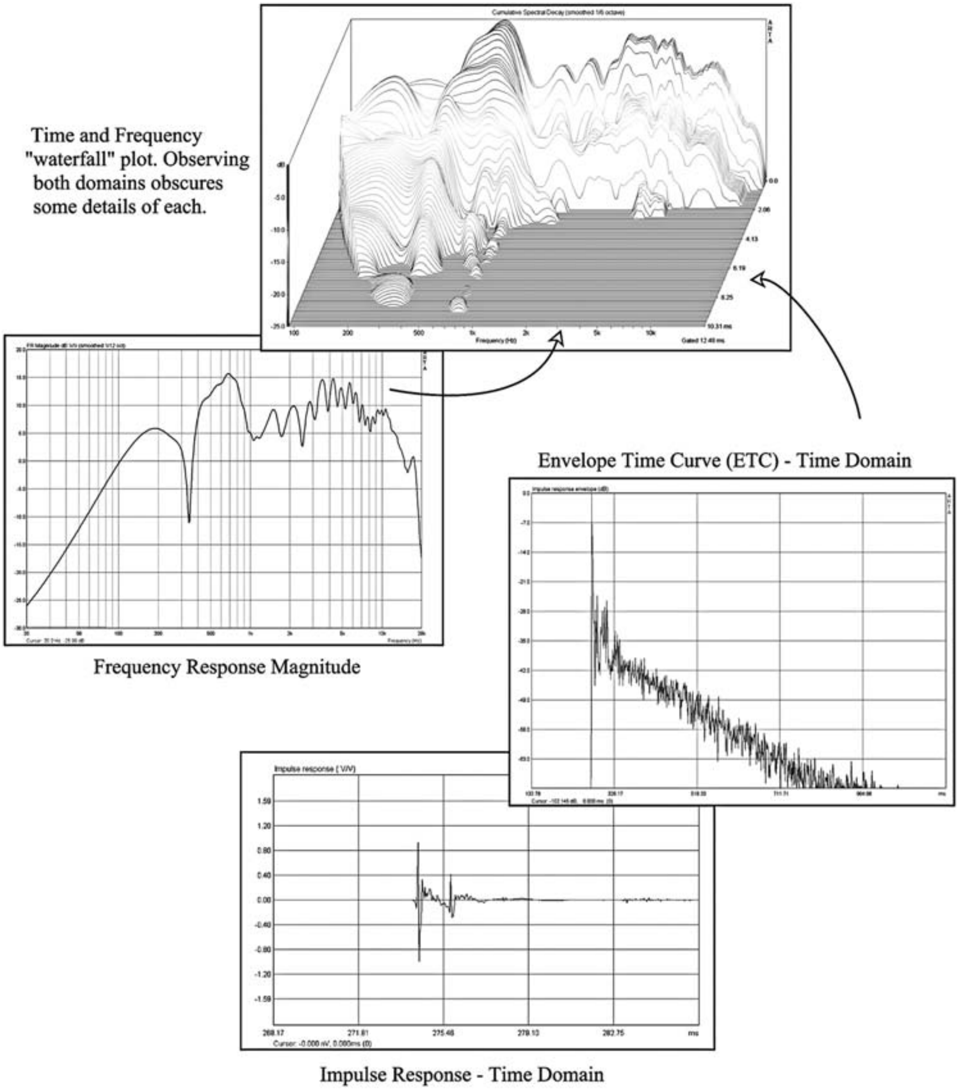


Figure 46-14. The FFT can be used to view the spectral content of a time domain measurement, Arta 1.2.

46.3.4.1 FFT Measurements

The Fourier Transform is a mathematical filtering process that determines the spectral content of a time domain signal. The Fast Fourier Transform, or FFT, is a computationally efficient version of the same. Most modern measurement systems make use of the computer's ability to quickly perform the FFT on sampled data. The cousin to the FFT is the IFFT, or Inverse Fast Fourier Transform. As one might guess, the IFFT takes a frequency domain signal as its input and produces a time domain signal. The FFT and IFFT form the bedrock of modern measurement systems. Many fields outside of audio use the FFT to analyze time records for periodic activity, such as utility companies to find peak usage times or an investment firm to investigate cyclic stock market behavior. Analyzers that use the Fast Fourier Transform to determine the spectral content of a time-varying signal are collectively called FFTs. If a broadband stimulus is used, the FFT can show the spectral response of the device under test (DUT). One such stimulus is the unit impulse, a signal of theoretically infinite amplitude and infinitely small time duration.

The FFT of such a stimulus is a straight, horizontal line in the frequency domain.

The time-honored hand clap test of a room is a crude but useful form of impulse response. The hand clap is useful for casual observations, but more accurate and repeatable methods are usually required for serious audio work. The drawbacks of using impulsive stimuli to measure a sound system include:

1. Impulses can drive loudspeakers into nonlinear behavior.
2. Impulse responses have poor signal-to-noise ratios, since all of the energy enters the system at one time and is reacquired over a longer span of time along with the noise from the environment.
3. There is no way to create a perfect impulse, so there will always be some uncertainty as to whether the response characteristic is that of the system, the impulse, or some nonlinearity arising from impulsing a loudspeaker.

Even with its drawbacks, impulse testing can provide useful information about the response of a loudspeaker or room.

46.3.4.2 Dual-Channel FFT

When used for acoustic measurements, dual-channel FFT analyzers digitally sample the signal fed to the loudspeaker, and also digitally sample the acoustic signal from the loudspeaker at the output of a test microphone. The signals are then compared by division, yielding the transfer function of the loudspeaker. Dual-channel FFTs have the advantage of being able to use any broadband stimulus as a test signal. This advantage is offset somewhat by poorer signal-to-noise performance and stability than other types of measurement systems, but the performance is often adequate for many measurement chores. Pink noise and swept sines provide much better stability and noise immunity. It is a computationally intense method since both the input and output signal must be measured simultaneously and compared, often in real time. For a proper comparison to yield a loudspeaker transfer function, it is important that the signals being compared have the same level, and that any time offsets between the two signals be removed. Dual-channel FFT analyzers have set up routines that simplify the establishment of these conditions. Portable computers have A/D converters as part of their on-board sound system, as well as a microprocessor to perform the FFT. With the appropriate software and sound system interface they form a powerful, low-cost and portable measurement platform.

46.3.4.3 Maximum-Length Sequence

The maximum-length sequence (MLS) is a pseudorandom noise test stimulus. The MLS overcomes some of the shortcomings of the dual-channel FFT, since it does not require that the input signal to the system be measured. A binary string (ones and zeros) is fed to the device under test while simultaneously being stored for future correlation with the loudspeaker response acquired by the test microphone. The pseudorandom sequence has a white spectrum (equal energy per Hz), and is exactly known and exactly repeatable. Comparing the input string with the string acquired by the test microphone yields the transfer function of the system. The advantage of the MLS is its excellent noise immunity and fast measurement time, making it a favorite of loudspeaker designers. A disadvantage is that the noise-like stimulus can be annoying, sometimes requiring that measurements be done after hours. The use of MLS has waned in recent years to log-swept sine measurements made on dual-channel FFT analyzers.

46.3.4.4 Time-Delay Spectrometry (TDS)

TDS is a fundamentally different method of measuring the transfer function of a system. Richard Heyser, a staff scientist at the Jet Propulsion Laboratories, invented the method. An anthology of Mr. Heyser's papers on TDS is available in the reference. Both the dual-channel FFT and MLS methods involve digital sampling of a broadband stimulus. TDS uses a method borrowed from the world of sonar, where a single-frequency sinusoidal "chirp" signal is fed to the system under test. The chirp slowly sweeps through the frequencies being measured, and is reacquired with a tracking filter by the TDS analyzer. The reacquired signal is then mixed with the outgoing signal, producing a series of sum and difference frequencies, each frequency corresponding to a different arrival time of sound at the microphone. The difference frequencies are transformed to the time domain with the appropriate transform, yielding the envelope-time Curve (ETC) of the system under test. TDS is based on the frequency domain, allowing the tracking filter to be tuned to the desired signal while ignoring signals outside of its bandwidth. TDS offers excellent noise immunity, allowing good data to be collected under near-impossible measurement conditions. Its downside is that good low-frequency resolution can be difficult to obtain without extended measurement times, plus the correct selection of measurement parameters requires a knowledgeable user. In spite of this, it is a favorite among contractors and consultants, who must often perform sound system calibrations in the real world of air conditioners, vacuum cleaners, and building occupants.

While other measurement methods exist, the ones outlined above make up the majority of methods used for field and lab testing of loudspeakers and rooms. Used properly, any of the methods can provide accurate and repeatable measured data. Many audio professionals have several measurement platforms and exploit the strong points of each when measuring a sound system.

46.3.5 Preparation

There are many measurements that can be performed on a sound system. A prerequisite to any measurement is to answer the following questions:

1. What am I trying to measure?
2. Why am I trying to measure it?
3. Is it audible?
4. Is it relevant?

Failure to consider these questions can lead to hours of wasted time and a hard drive full of meaningless data. Even with the incredible technologies that we have available to us, the first part of any measurement session is to listen. It can take many hours to determine what needs to be measured to solve a sound system problem, yet the actual measurement itself can often be completed in seconds. Using an analogy from the medical field, the physician must query the patient at length to narrow down the ailment. The more that is known about the ailment, the more specific and relevant the tests that can be run for diagnosis. There is no need to test for tonsillitis if the problem is a sore back!

1. What am I measuring? A fundamental decision that precedes a meaningful measurement is how much of the room's response to include in the measured data. Modern measurement systems have the ability to perform semianechoic measurements, and the measurer must decide if the loudspeaker, the room, or the combination needs to be measured. If one is diagnosing loudspeaker ailments, there is little reason to select a time window long enough to include the effects of late reflections and reverberation. A properly selected time window can isolate the direct field of the loudspeaker and allow its response to be evaluated independently of the room. If one is trying to measure the total decay time of the room, the direct sound field becomes less important, and a microphone placement and time window are selected to capture the entire energy decay. Most modern measurement systems acquire the complete impulse response, including the room decay, so the choice of the time window size can be made after the fact during post processing.
2. Why am I measuring? There are several reasons for performing acoustic measurements in a space. An important reason for the system designer is to characterize the listening environment. Is it dead? Is it live? Is it reverberant? These questions must be considered prior to the design of a sound system for the space. While the human hearing system can provide the answers to these questions, it cannot document them and it is easily deceived. Measurements might also be performed to document the performance of an existing system prior to performing changes or adding room treatment. Customers sometimes forget how bad it once sounded after a new or upgraded system is in place for a few weeks.

   The most common reason for performing measurements on a system is for calibration purposes. This can include equalization, signal alignment, crossover selection, and a multiplicity of other reasons. Since loudspeakers interact in a complex way with their environment, the final phase of any system installation is to verify system performance by measurement.
3. Is it audible? Can I hear what I am trying to measure? If one cannot hear an anomaly, there is little reason to attempt to measure it. The human hearing system is perhaps the best tool available for determining what should be measured about a sound system. The human hearing system can tell us that something doesn't sound right, but the cause of the problem can be revealed by measurement. Anything you can hear can be measured, and once it is measured it can be quantified and manipulated.
4. Is it relevant? Am I measuring something that is worth measuring? If one is working for a client, time is money. Measurements must be prioritized to focus on audible problems. Endless hours can be spent "chasing rabbits" by measuring details that are of no importance to the client. This is not necessarily a fruitless process, but it is one that should be done on your own time. I have on several occasions spent time measuring and documenting anomalies that had nothing to do with the customer's reason for calling me. All venues have problems that the owner is unaware of. Communication with the client is the best way to avoid this pitfall.

46.3.5.1 Dissecting the Impulse Response

The audio practitioner is often faced with the dilemma of determining whether the reason for bad sound is the loudspeaker system, the room, or an interaction of the two. The impulse response can hold that answer to these and other perplexing questions. The impulse response in its amplitude versus time display is not particularly useful for other than determining the polarity of a system component, Fig. 46-15. A better representation comes from squaring impulse response (making all deflections positive) and displaying the square root of the result on a logarithmic vertical scale. This log-squared response allows the relative levels of energy arrivals to be compared, Fig. 46-16.

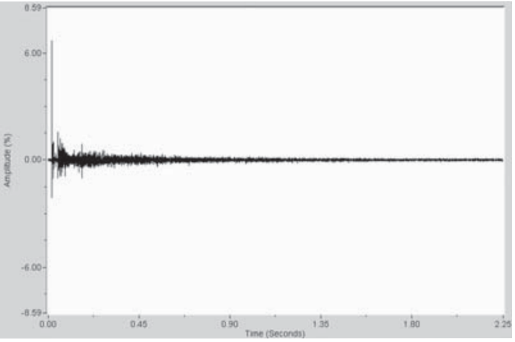


Figure 46-15. The impulse response, SIA-SMAART.

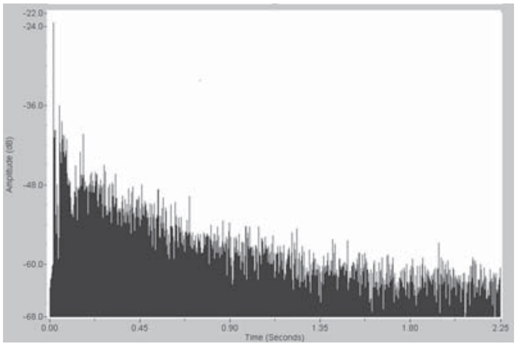


Figure 46-16. The log-squared response, SIA-SMAART.

46.3.5.2 The Envelope-Time Curve

Another useful way of viewing the impulse response is in the form of the envelope-time curve, or ETC. The ETC is also a contribution of Richard Heyser.[2] It takes the real part of the impulse response and combines it with a 90 degrees phase shifted version of the same, Fig. 46-17. One way to get the shifted version is to use the Hilbert Transform. The complex combination of these two signals yields a time domain waveform that is often easier to interpret than the impulse response. The ETC can be loosely thought of as a smoothing function for the log-squared response, showing the envelope of the data. This can be more revealing as to the audibility of an event. The impulse response, log-squared response, and energy-time curve are all different ways to view the time domain data.

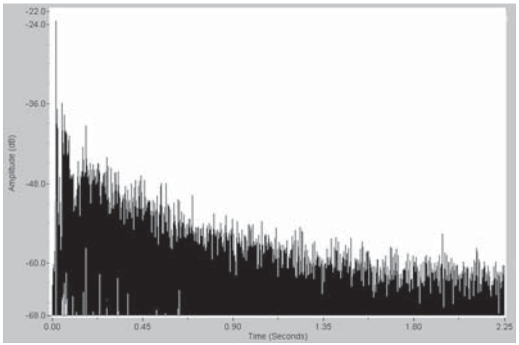


Figure 46-17. The envelope-time curve (ETC), SIA-SMAART.

46.3.5.3 A Global Look

When starting a measurement session, a practical approach is to first take a global look and measure the complete decay of the room. The measurer can then choose to ignore part of the time record by using a time window to isolate the desired part during postprocessing. The length of the time window can be increased to include the effects of more of the energy returned by the room. The time window can also be used to isolate a reflection and view its spectral content. Just like your life span represents a time window in human history, a time window can be used to isolate parts of the impulse response.

46.3.5.4 Time Window Length

The time domain response can be divided to identify the portion that can be attributed to the loudspeaker and that which can be attributed to the room. It must be emphasized that there is a rather gray and frequency-dependent line between the two, but for this discussion we will assume that we can clearly separate them. The direct field is the energy that arrives at the listener prior to any reflections from the room. The division is fairly distinct if neither the loudspeaker nor microphone is placed near any reflecting surfaces, which, by the way, is a good system design practice. At long wavelengths (low frequencies) the direct field may include the effects of boundaries near the loudspeaker and microphone. As frequency increases, the sound from the loudspeaker becomes less affected by boundary effects (due in part to increased directivity) and can be measured independently of them. Proper loudspeaker placement produces a time gap between the sound energy arrivals from the loudspeaker and the later arriving room response. We can use this time gap to aid in selecting a time window to separate the loudspeaker response from the room response and diagnosing system problems.

46.3.5.5 Acoustic Wavelengths

Sound travels in waves. The sound waves that we are interested in characterizing have a physical size. There will be a minimum time span required to observe the spectral response of a waveform. The minimum required length of time to view an acoustical event is determined by the longest wavelength (lowest frequency) present in the event. At the upper limits of human hearing, the wavelengths are only a few millimeters in length, but as frequency decreases the waves become increasingly larger. At the lowest frequencies that humans hear, the wavelengths are many meters long, and can actually be larger than the listening (or measurement) space. This makes it difficult to measure low frequencies from a loudspeaker independently of the listening space, since low frequencies radiated from a loudspeaker interact (couple) with the surfaces around them. In an ideally positioned loudspeaker, the first energy arrival from the loudspeaker at mid- and high frequencies has already dissipated prior to the arrival of reflections and can therefore often be measured independently of them. The human hearing system tends to fuse the direct sound from the loudspeaker with the early reflections from nearby surfaces with regard to level (loudness) and frequency (tone). It is usually useful to consider them as separate events, especially since the time offset between the direct sound and first reflections will be unique for each listening position. This precludes any type of frequency domain correction (i.e., equalization) of the room/loudspeaker response other than at frequencies where coupling occurs due to close proximity to nearby surfaces. While it is possible to compensate to some extent for room reflections at a point in space (acoustic echo cancellers used for conference systems), this correction cannot be extended to include an area. This inability to compensate for the reflected energy at mid/high frequencies suggests that their effects be removed from the loudspeaker's direct field response prior to meaningful equalization work by use of an appropriate time window.

46.3.5.6 Microphone Placement

A microphone is needed to acquire the sound radiated into the space from the loudspeaker at a discrete position. Proper microphone placement is determined by the type of test being performed. If one were interested in measuring the decay time of the room, it is usually best to place the microphone well beyond critical distance. This allows the build-up of the reverberant field to be observed as well as providing good resolution of the decaying tail. Critical distance is the distance from the loudspeaker at which the direct field level and reverberant field level are equal. It is described further in Section 46.3.5.7. If it's the loudspeaker's response that needs to be measured, then a microphone placement inside of critical distance will provide better data on some types of analyzers, since the direct sound field is stronger relative to the later energy returning from the room. If the microphone is placed too close to the loudspeaker, the measured sound levels will be accurate for that position, but may not accurately extrapolate to greater distances with the inverse-square law. As the sound travels farther, the response at a remote listening position may bear little resemblance to the response at the near field microphone position. For this reason, it is usually desirable to place the microphone in the far free field of the loudspeaker—not too close and not too far away. The approximate extent of the near field can be determined by considering that the path length difference from the measurement position (assumed axial) and the edge of the sound radiator should be less than $^1/_4$ wavelength at the frequency of interest. This condition is easily met for a small loudspeaker that is radiating low frequencies. Such devices closely approximate an ideal point source. As the frequency increases the condition becomes more difficult to satisfy, especially if the size of the radiator also increases. Large radiators (or groups of radiators) emitting high frequencies can extend the near field to very long distances. Line arrays make use of this principle to overcome the inverse-square law. In practice, small bookshelf loudspeakers can be accurately measured at a few meters. About 10 m is a common measurement distance for moderate-sized, full-range loudspeakers in a large space. Even greater distances are required for large devices radiating high frequencies. A general guideline is to not put the mic closer than three times the loudspeaker's longest dimension.

46.3.5.7 Estimate the Critical Distance D_C

Critical distance is easy to estimate. A quick method with adequate accuracy requires a sound level meter and noise source. Ideally, the noise source should be band limited, as critical distance is frequency dependent. The 2 kHz octave band is a good place to start when measuring critical distance. Proceed as follows:

1. Energize the room with pink noise in the desired octave band from the sound source being measured. The level should be at least 25 dB higher than the background noise in the same octave band.
2. Using the sound level meter, take a reading near the loudspeaker (about 1 m) and on-axis. At this distance, the direct sound field will dominate the measurement.
3. Move away from the loudspeaker while observing the sound level meter. The sound level will fall off as you move farther away. If you are in a room with a reverberant sound field, at some distance the meter reading will quit dropping. You have now moved beyond critical distance. Measurements of the direct field beyond this point will be a challenge for some types of analysis. Move back toward the loudspeaker until the meter begins to rise again. You are now entering a good region to perform acoustic measurements on loudspeakers in this environment. The above process provides an estimate that is adequate for positioning a measurement microphone for loudspeaker testing. With a mic placement inside of critical distance, the direct field is a more dominant feature on the impulse response and a time window will be more effective in removing room reflections.

At this point it is interesting to wander around the room with the sound level meter and evaluate the uniformity of the reverberant field. Rooms that are reverberant by the classical definition will vary little in sound level beyond critical distance when energized with a continuous noise spectrum. Such spaces have low internal sound absorption relative to their volume.

46.3.5.8 Common Factors to All Measurement Systems

Let's assume that we wish to measure the impulse response of a loudspeaker/room combination. While it would not be practical to measure the response at every seat, it is good measurement practice to measure at as many seats as are required to prove the performance of the system. Once the impulse response is properly acquired, any number of postprocesses can be performed on the data to extract information from it. Most modern measurement systems make use of digital sampling in acquiring the response of the system. The fundamentals and prerequisites are not unlike the techniques used to make any digital recording, where one must be concerned with the level of an event and its time length. Some setup is required and some fundamentals are as follows:

1. The sampling rate must be fast enough to capture the highest frequency component of interest. This requires at least two samples of the highest frequency component. If one wished to measure to 20 kHz, the required sample rate would need to be at least 40 kHz. Most measurement systems sample at 44.1 kHz or 48 kHz, more than sufficient for acoustic measurements.
2. The time length of the measurement must be long enough to allow the decaying energy curve to flatten out into the room noise floor. Care must be taken to not cut off the decaying energy, as this will result in artifacts in the data, like a scratch on a phonograph record. If the sampling rate is 44.1 kHz, then 44,100 samples must be collected for each second of room decay. A 3-second room would therefore require 44.1 × 1000 × 3 or 128,000 samples. A hand clap test is a good way to estimate the decay time of the room and therefore the required number of samples to fully capture it. The time span of the measurement also determines the lowest frequency that can be resolved from the measured data, which is approximately the inverse of the measurement length. The sampling rate can be reduced to increase the sampling time to yield better low-frequency information. The trade-off is a reduction in the highest frequency that can be measured, since the condition outlined in step one may have been violated.
3. The measurement must hav a sufficient signal-to-noise ratio to allow the decaying tail to be fully observed. This often requires that the measurement be repeated a number of times and the results averaged. Using a dual-channel FFT or MLS, the improvement in SNR will be 3 dB for each doubling of the number of averages. Ten averages is a good place to start, and this number can be increased or decreased depending on the environment. The level of the test stimulus is also important. Higher levels produce improved SNR, but can also stress the loudspeaker.
4. Perform the test and observe the data. It should fill the screen from top left to bottom right and be fully

decayed prior to reaching the right side of the screen. It should also be repeatable. Run the test several times to check for consistency. Background noise can dramatically affect the repeatability of the measurement and the validity of the data.

Once the impulse response is acquired, it can be further analyzed for spectral content, intelligibility information, decay time, etc. These are referred to as *metrics*, and some require some knowledge on the part of the measurer in properly placing markers (called *cursors*) to identify the parameters required to perform the calculations. Let us look at how the response of the loudspeaker might be extracted from the data just gathered.

The time domain data displays what would have resulted if an impulse were fed through the system. Don't try to correlate what you see on the analyzer with what you heard during the test. Most measurement systems display an impulse response that is calculated from a knowledge of the input and output signal to the system, and there is no resemblance between what you hear when the test is run and what you are seeing on the screen, Fig. 46-18.

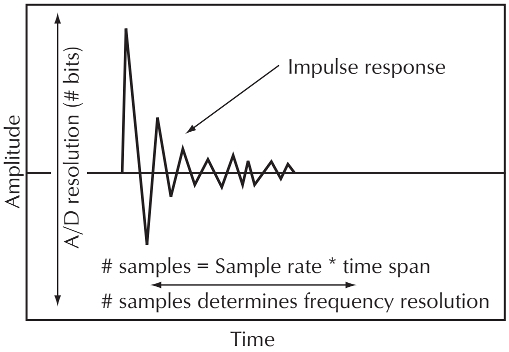

Amplitude
A/D resolution (# bits)
Impulse response
# samples = Sample rate * time span
# samples determines frequency resolution
Time

Figure 46-18. Many analyzers acquire the room response by digital sampling.

We can usually assume that the first energy arrival is from the loudspeaker itself, since any reflection would have to arrive later than the first wave front since it had to travel farther. Pre-arrivals can be caused by the acoustic wave propagating through a solid object, such as a ceiling or floor and reradiating near the microphone. Such arrivals are very rare and usually quite low in level. In some cases a reflection may actually be louder than the direct arrival. This could be due to loudspeaker design or its placement relative to the mic location. It's up to the measurer to determine if this is normal for a given loudspeaker position/seating position. All loudspeakers will have some internal and external reflections that will arrive just after the first wave front. These are actually a part of the loudspeaker's response and can't be separated from the first wave front with a time window due to their close proximity without extreme compromises in frequency resolution. Such reflections are at least partially responsible for the characteristic sound of a loudspeaker. Studio monitor designers and studio control room designers go to great lengths to reduce the level of such reflections, yielding more accurate sound reproduction. Good system design practice is to place loudspeakers as far as possible from boundaries (at least at mid- and high frequencies). This will produce an initial time gap between the loudspeaker's response and the first reflections from the room. This gap is a good initial dividing point between the loudspeaker's response and the room's response, with the energy to the left of the dividing cursor being the response of the loudspeaker and the energy to the right the response of the room. The placement of this divider can form a time window by having the analyzer ignore everything later in time than the cursor setting. The time window size also determines the frequency resolution of the post-processed data. In the frequency domain, improved resolution means a smaller number. For instance, 10 Hz resolution is better than 40 Hz resolution. Since time and frequency have an inverse relationship, the time window length required to observe 10 Hz will be much longer than the time window length required to resolve 40 Hz. The resolution can be estimated by $f = 1/T$, where T is the length of the time window in seconds. Since a frequency magnitude plot is made up of a number of data points connected by a line, another way to view the frequency resolution is that it is the number of Hz between the data points in a frequency domain display.

The method of determination of the time window length varies with different analyzers. Some allow a cursor to be placed anywhere on the data record, and the placement determines the frequency resolution of the spectrum determined by the window length. Others require that the measurer select the number of samples to be used to form the time window, which in turn determines the frequency resolution of the time window. The window can then be positioned at different places on the time domain plot to observe the spectral content of the energy within the window, Figs. 46-19, 46-20, and 46-21.

For instance, a 1 second total time (44,100 samples) could be divided into about twenty two time windows of 2048 samples each (about 45 ms). Each window would allow the observation of the spectral content down to ($\frac{1}{45}$) × 1000 or 22 Hz. The windows can be overlapped and moved around to allow more precise selection of the time span to be observed. Displaying a

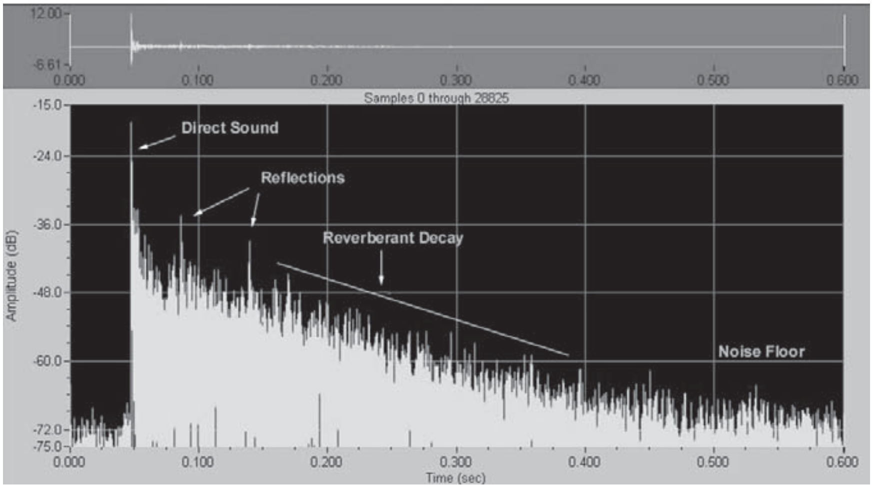


Figure 46-19. A room response showing the various sound fields that can exist in an enclosed space, SIA-SMAART.

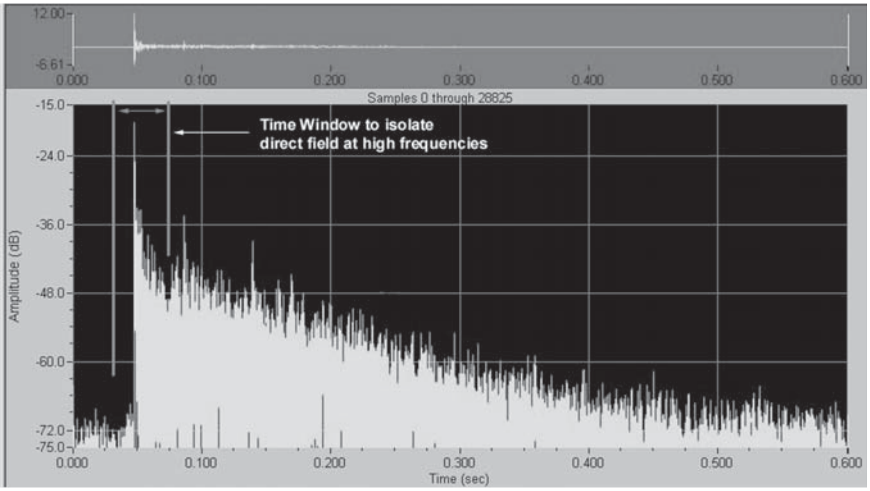


Figure 46-20. A time window can be used to isolate the loudspeaker's response from the room reflections.

number of these time windows in succession, each separated by a time offset, can form a 3D plot known as a waterfall.

46.3.5.9 Data Windows

There are some conditions that must be observed when placing cursors to define the time window. Ideally, we would like to place the cursor at a point on the time record where the energy is zero. A cursor placement that cuts off an energy arrival will produce a sharp rise or fall time that produces artifacts in the resultant calculated spectral response. Discontinuities in the time domain have broad spectral content in the frequency domain. A good example is a scratch on a phonograph record. The discontinuity formed by the scratch manifests itself as a broadband click during playback. If an otherwise smooth wheel has a discontinuity at one point, it would thump annoyingly when it was rolled on a smooth surface. Our measurement systems treat the data within the selected window as a continuously repeating event. The end of the event must line up with the beginning or a discontinuity occurs resulting in the generation of high-frequency artifacts called *spectral leakage*. In the same manner that a physical discontinuity in a phonograph record or wheel can be corrected by polishing, a discontinuity in a sampled time measure-

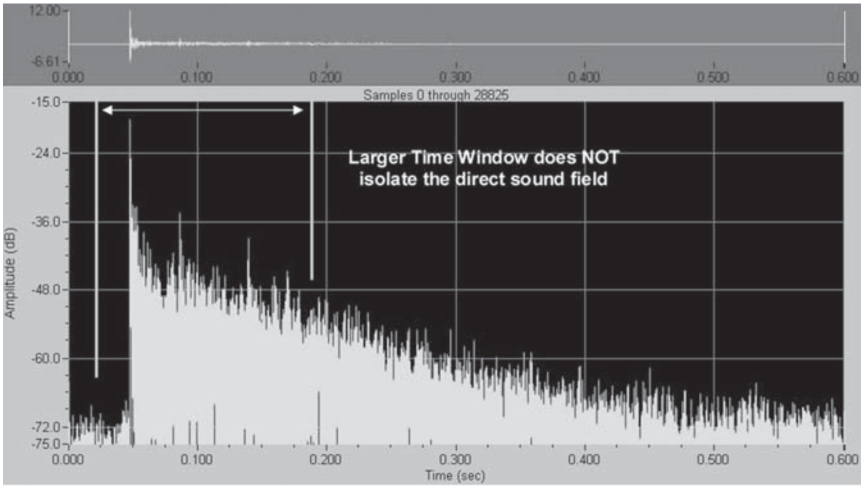


Figure 46-21. Increasing the length of the time window increases the frequency resolution, but lets more of the room into the measurement, SIA-SMAART.

ment can be remedied by tapering the energy at the beginning and end of the window to zero using a mathematical function. A number of data window shapes are available for performing the smoothing.

These include the Hann, Hamming, Blackman-Harris, and others. In the same way that a physical polishing process removes some good material from what is being rubbed, data windows remove some good data in the process of smoothing the discontinuity. Each window has a particular shape that leaves the data largely untouched at the center of the window but tapers it to varying degrees toward the edges. Half windows only smooth the data at the right edge of the time record while full windows taper both (start and stop) edges. Since all windows have side effects, there is no clear preference as to which one should be used. The Hann window provides a good compromise between time record truncation and data preservation. Figs. 46-22 and 46-23 show how a data window might be used to reduce spectral leakage.

46.3.5.10 A Methodical Approach

Since there are an innumerable number of tests that can be performed on a system, it makes sense to establish a methodical and logical process for the measurement session. One such scenario may be as follows:

1. Determine the reason for and scope of the measurement session. What are you looking for? Can you hear it? Is it repeatable? Why do you need this information?
2. Determine what you are going to measure. Are you looking at the room or at the sound system? If it is the room, possibly the only meaningful measurements will be the overall decay time and the noise floor. If you are looking at the sound system, decide if you need to switch off or disconnect some loudspeakers. This may be essential to determine whether the individual components are working properly, or that an anomaly is the result of interaction between several components. "Divide and conquer" is the axiom.

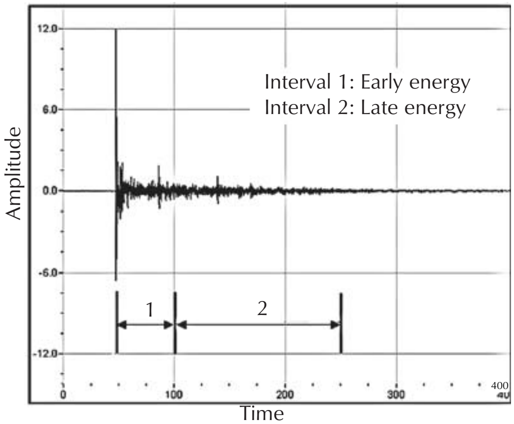


Figure 46-22. The impulse response showing both early and late energy arrivals.

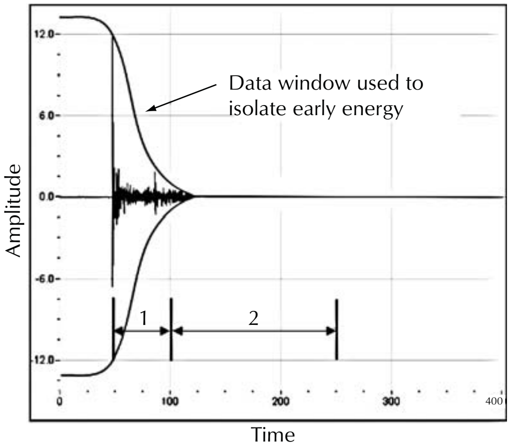


Figure 46-23. A data window is used to remove the effects of the later arrivals.

3. Select the microphone position. I usually begin by looking at the on-axis response of the loudspeaker as measured from inside of critical distance. If multiple loudspeakers are on, turn all of them off but one prior to measuring. The microphone should be placed in the far free field of the loudspeaker as previously described. When measuring a loudspeaker's response, care should be taken to eliminate the effects of early reflections on the measured data, as these will generate acoustic comb filters that can mask the true response of the loudspeaker. In most cases the predominant offending surface will be the floor or other boundaries near the microphone and loudspeaker. These reflections can be reduced or eliminated by using a ground plane microphone placement, a tall microphone stand (when the loudspeaker is overhead), or some strategically placed absorption. I prefer the tall microphone stand for measuring installed systems with seating present since it works most anywhere, regardless of the seating type. The idea is to intercept the sound on its way to a listener position, but before it can interact with the physical boundaries around that position. These will always be unique to that particular seat, so it is better to look at the free field response, as it is the common denominator to many listener seats.
4. Begin with the big picture. Measure an impulse response of the complete decay of the space. This yields an idea of the overall properties of the room/system and provides a good point of reference for zooming in to smaller time windows. Save this information for documentation purposes, as later you may wish to reopen the file for further processing.
5. Reduce the size of the time window to eliminate room reflections. Remember that you are trading off frequency resolution when truncating the time record, Fig. 46-24. Be certain to maintain sufficient resolution to allow adequate low-frequency detail. In some cases, it may be impossible to maintain a sufficiently long window to view low frequencies and at the same time eliminate the effects of reflections at higher frequencies, Fig. 46-25. In such cases, the investigator may wish to use a short window for looking at the high-frequency direct field, but a longer window for evaluating the woofer. Windows appropriate for each part of the spectrum can be used. Some measurement systems provide variable time windows, which allow low frequencies to be viewed in great detail (long time window) while still providing a semianechoic view (short time window) at high frequencies. There is evidence to support that this is how humans process sound information, making this method particularly interesting, Fig. 46-26.
6. Are other microphone positions necessary to characterize this loudspeaker? The off-axis response of some loudspeakers is very similar to the on-axis response, reducing the need to measure at many angles. Other loudspeakers have very erratic responses, and a measurement at any one point around the loudspeaker may bear little resemblance to the response at other positions. This is a design issue, but one that must be considered by the measurer.
7. Once an accurate impulse response is measured, it can be postprocessed to yield information on spectral content, speech intelligibility, and music clarity. There are a number of metrics that can provide this information. These are interpretations of the measured data and generally correlate with subjective perception of the sound at that seat.
8. An often overlooked method of evaluating the impulse response is the use of convolution to encode it onto anechoic program material. An excellent freeware convolver called GratisVolver is available from www.catt.se. Listening to the IR can often reveal subtleties missed by the various metrics, as well as provide clues as to what postprocess must be used to observe the event of interest.

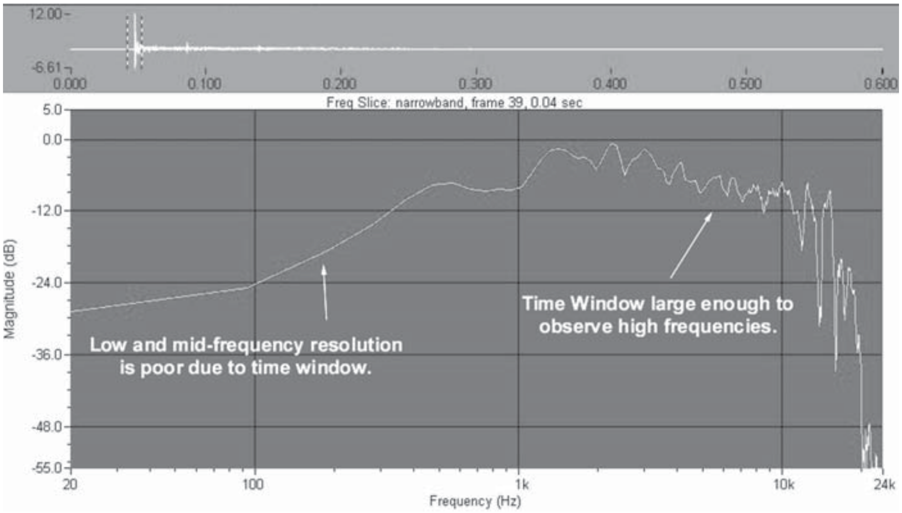


Figure 46-24. A short time window isolates the direct field at high frequencies at the expense of low-frequency resolution, SIA-SMAART.

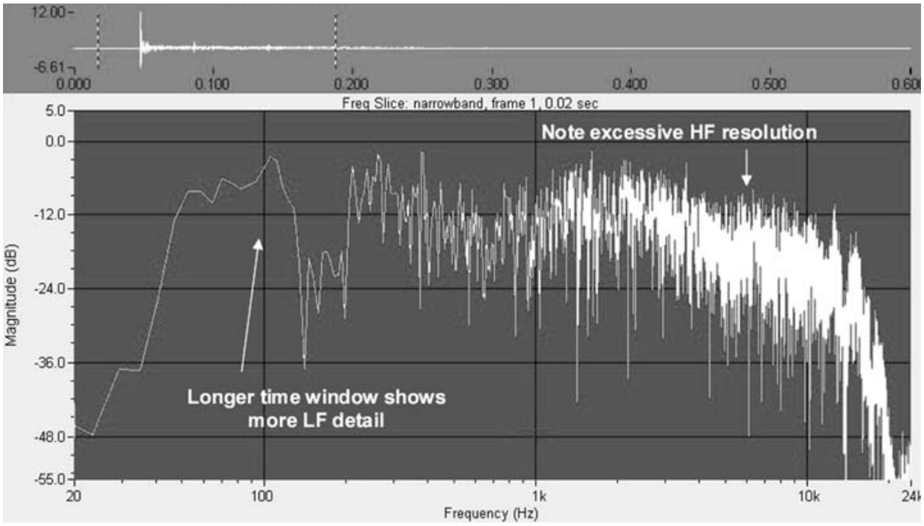


Figure 46-25. A long time window provides good low-frequency detail, SIA-SMAART.

46.3.6 Human Perception

Useful measurement systems can measure the impulse response of a loudspeaker/room combination with great detail. Information regarding speech intelligibility and music clarity can be derived from the impulse response. In nearly all cases, this involves postprocessing the impulse response using one of several clarity measure metrics.

46.3.6.1 Percentage Articulation Loss of Consonants-(%Alcons)

For speech, one such metric is the percentage articulation loss of consonants, or *%Alcons*. Though not in widespread use today, a look at it can provide insight into the requirements for good speech intelligibility. A *%Alcons* measurement begins with an impulse response, which is usually displayed as a log-squared response or ETC. Since the calculation essentially

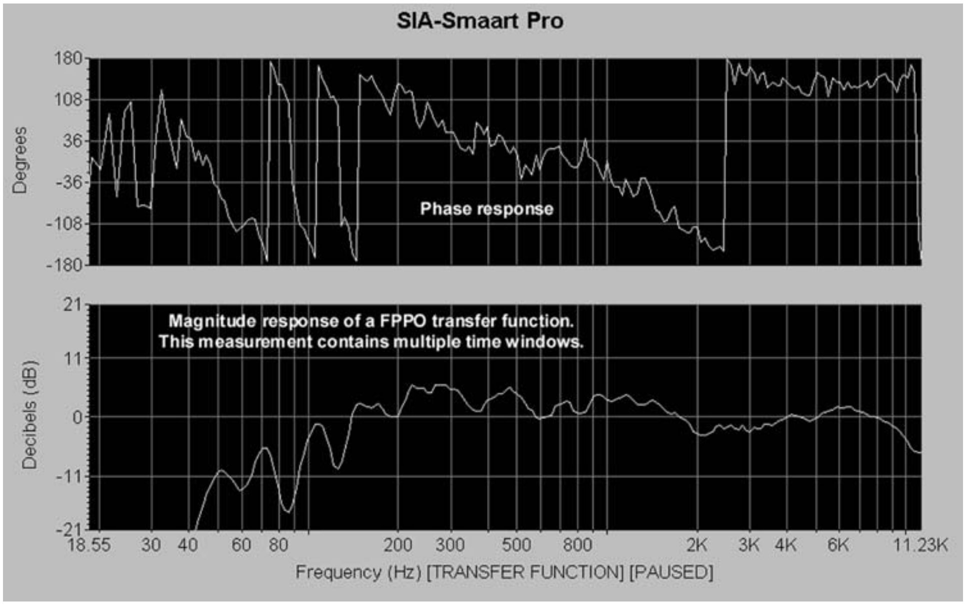
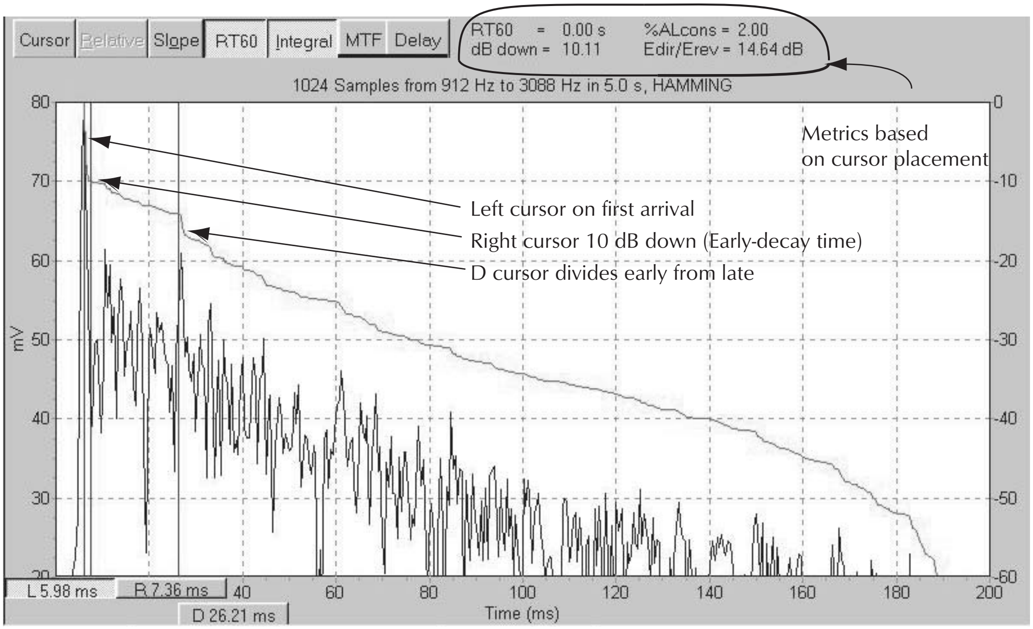


Figure 46-26. The time window increases with length as frequency decreases.


Cursor
Relative
Slope
RT60
Integral
MTF
Delay
1024 Samples from 912 Hz to 3088 Hz in 5.0 s, HAMMING
Metrics based
on cursor placement
Left cursor on first arrival
Right cursor 10 dB down (Early-decay time)
D cursor divides early from late
mV
80
70
60
50
40
30
20
0
-10
-20
-30
-40
-50
-60
40
60
80
100
120
140
160
180
200
Time (ms)

RT60 = 0.00 s
dB down = 10.11
%ALcons = 2.00
Edir/Erev = 14.64 dB

L 5.98 ms
R 7.36 ms
D 26.21 ms

Figure 46-27. The ETC can be processed to yield an intelligibility score, TEF25.

examines the ratio between early energy, late energy, and noise, the measurer must place cursors on the display to define these parameters. These cursors may be placed automatically by the measurement program. The result is weighted with regard to decay time, so this too must be defined by the measurer. Analyzers such as the TEF25™ and EASERA include best guess default placements based on the research of Peutz, Davis, and others, Fig. 46-27.

These placements were determined by correlating measured data with live listener scores in various acoustic environments, and represent a defined and orderly approach to achieving meaningful results that correlate with the perception of live listeners. The measurer is free to choose alternate cursor placements, but great care must be taken to be consistent. Also, alternate cursor placements make it difficult if not impossible to compare your results with those obtained by other measurers. In the default %Alcons placement, the early energy (direct sound field) includes the first major sound arrival and any energy arrivals within the next 7–10 ms. This forms a tight time span for the direct sound. Energy beyond this span is considered late energy and an impairment to communication. As one might guess, a later cursor placement yields better intelligibility scores, since more of the room response is being considered beneficial to intelligibility. As such, the default placement yields a worst-case scenario. The default placement considers the effects of the early-decay time (EDT) rather than the classical T_{30} since short EDTs can yield good intelligibility, even in rooms with a long T_{30}. Again, the measurer is free to select an alternative cursor placement for determining the decay time used in the calculation, with the same caveats as placing the early-to-late dividing cursor. The *%Alcons* score is displayed instantly upon cursor placement and updates as the cursors are moved.

46.3.6.2 Speech Transmission Index—(STI)

The STI can be calculated from the measured impulse response with a routine outlined by Schroeder and detailed by Becker in the reference. The STI is probably the most widely used contemporary measure of intelligibility. It is supported by virtually all measurement platforms, and some handheld analyzers are available for quick checks. In short, it is a number ranging from 0 to 1, with fair intelligibility centered at 0.5 on the scale. For more details on the Speech Transmission Index, see the chapter on speech intelligibility in this text.

46.3.7 Polarity

Good sound system installation practice dictates maintaining proper signal polarity from system input to system output. An audio signal waveform always swings above and below some reference point. In acoustics, this reference point is the ambient atmospheric pressure. In an electronic device, the reference is the 0 VA reference of the power supply (often called *signal ground*) in push-pull circuits or a fixed dc offset in class A circuits. Let's look at the acoustic situation first. An increase in the air pressure caused by a sound wave will produce an inward deflection of the diaphragm of a pressure microphone (the most common type) regardless of the microphone's orientation toward the source. This inward deflection should cause a positive-going voltage swing at the output of the microphone on pin 2 relative to pin 3, as well as at the output of each piece of equipment that the signal passes through. Ultimately the electrical signal will be applied to a loudspeaker, which should deflect outward (toward an axial listener) on the positive-going signal, producing an increase in the ambient atmospheric pressure. Think of the microphone diaphragm and loudspeaker diaphragm moving in tandem and you will have the picture. Since most sound reinforcement equipment uses bipolar power supplies (allowing the audio signal to swing positive and negative about a zero reference point), it is possible for signals to become inverted in polarity (flipped over). This causes a device to output a negative-going voltage when it is fed a positive-going voltage. If the loudspeaker is reverse-polarity from the microphone, an increase in sound pressure at the microphone (compression) will cause a decrease in pressure in front of the loudspeaker (rarefaction). Under some conditions, this can be extremely audible and destructive to sound quality. In other scenarios it can be irrelevant, but it is always good to check.

System installers should always check for proper polarity when installing the sound system. There are a number of methods, some simple and some complex. Let's deal with them in order of complexity, starting with the simplest and least-costly method.

46.3.7.1 The Battery Test

Low-frequency loudspeakers can be tested using a standard 9 V battery. The battery has a positive and negative terminal, and the spacing between the terminals is just about right to fit across the terminals of most woofers. The loudspeaker cone will move outward when the battery is placed across the loudspeaker terminals with the battery positive connected to the loudspeaker positive. While this is one of the most accurate methods for testing polarity, it doesn't work for most electronic devices or high-frequency drivers. Even so, it's probably the least-costly and most accurate way to test a woofer.

46.3.7.2 Polarity Testers

There are a number of commercially available polarity test sets in the audio marketplace. The set includes a sending device that outputs a test pulse, Fig. 46-28, through a small loudspeaker (for testing microphones) or an XLR connector (for testing electronic devices) and a receiving device that collects the signal via an internal microphone (loudspeaker testing) or XLR input jack. A green light indicates correct polarity and a red light indicates reverse polarity. The receive unit should be placed at the system output (in front of the loudspeaker) while the send unit is systematically moved from device to device toward the system input. A polarity reversal will manifest itself by a red light on the receive unit.

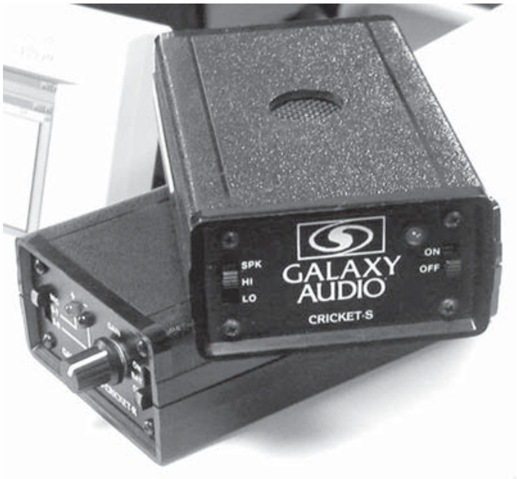


Figure 46-28. A popular polarity test set.

46.3.7.3 Impulse Response Tests

The impulse response is perhaps the most fundamental of audio and acoustic measurements. The polarity of a loudspeaker or electronic device can be determined from observing its impulse response, Figs. 46-29 and 46-30. This is one of the few ways to test flown loudspeakers from a remote position. It is best to test the polarity of components of multiway loudspeakers individually, since all of the individual components may not be polarized the same. Filters in the signal path (i.e., active crossover network) make the results more difficult to interpret, so it may be necessary to carefully test a system component (i.e., woofer) full-range for definitive results. Be sure to return the crossover to its proper setting before continuing.

46.4 Conclusion

The test and measurement of the sound reinforcement system are a vital part of the installation and diagnostic processes. The FFT and the analyzers that use it have revolutionized the measurement process, allowing sound practitioners to pick apart the system response and look at the response of the loudspeaker, roo, or both. Powerful analyzers that were once beyond the reach of most technicians are readily available and affordable, and cost can no longer be used as an excuse for not measuring the system. The greatest investment by far is the time required to grasp the fundamentals of acoustics to allow interpretation of the data. Some of this information is general, and some of it is specific to certain measurement systems.

The acquisition of a measurement system is the first step in ascending the capability and credibility ladder. The next steps include acquiring proper instruction on its use by self-study or short course. The final and most important steps are the countless hours in the field required to correlate measured data with the hearing process. As proficiency in this area increases, the speed of execution, validity, and relevance of the measurements will increase also. While we can all learn how to make the measurements in a relatively short time span, the rest of our careers will be spent learning how to interpret what we are measuring.

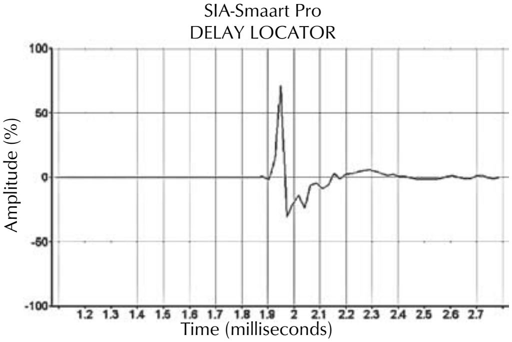


Figure 46-29. The impulse response of a transducer with correct polarity.

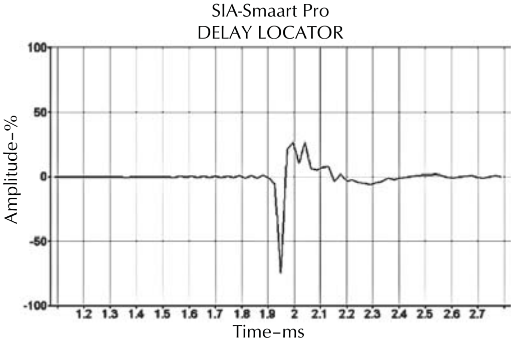


Figure 46-30. The impulse response of a reverse-polarity transducer.

References

1. D. Davis and C. Davis, *Sound System Engineering*, Focal Press, 1997, Boston, MA.
2. The Heyser Anthology, Audio Engineering Society, NY, NY.

Bibliography

Understanding Sound System/Room Interactions, Sam Berkow, Syn-Aud-Con Tech Topic, Volume 28, Number 2.

Joseph D'Appolitto, *Testing Loudspeakers*, Old Colony Sound Labs, Petersborough, NH.

This page intentionally left blank

Chapter 47

What's the Ear For? How to Protect It

by Les Blomberg and Noland Lewis

47.1 What's the Ear For? 1633
  47.1.1 What Does Hearing Damage Sound Like? 1633
47.2 How Loud Is Too Loud? OSHA, NIOSH, EPA, WHO 1633
47.3 Indicators of Hearing Damage 1635
47.4 Protecting Your Hearing 1635
  47.4.1 Protecting Concert-Goers and Other Listeners 1636
  47.4.2 Protecting the Community. 1636
47.5 Too Much of a Good Thing 1637
  47.5.1 A Compliance and Enforcement Tool. 1637
    47.5.1.1 The SLARMSolution™ 1637
    47.5.1.2 SLARMSoft™ Software Suite 1638
    47.5.1.3 SLARM™ Operation 1639
47.5.1.3.1 Applications 1641
47.6 Summary 1642

This page intentionally left blank

47.1 What's the Ear For?

An ear is for listening, and for the lucky few, listening to music is their job. But an ear is for much more—lose your hearing, and besides not hearing music, you lose your connection with other people. Hearing is the sense most related to learning and communication, and is the sense that connects you to ideas and other people. Helen Keller, who lost both her sight and hearing at a young age, said that hearing loss was the greater affliction for this reason.

To professionals in the music industry, their hearing is their livelihood. To be able to hear well is the basis for sound work. Protecting your hearing will determine whether you are still working in the industry when you are 64, or even whether you can still enjoy music, and it will determine whether you will hear your spouse and grandchildren then, too.

47.1.1 What Does Hearing Damage Sound Like?

Hearing loss is the most common preventable workplace injury. Ten million Americans have noise-induced hearing loss. Ears can be easily damaged, resulting in partial or complete deafness or persistent ringing in the ears.

Hearing loss isn't necessarily quiet. It can be a maddening, aggravating buzz or ringing in the ear, called *tinnitus*. Or it may result in a loss of hearing ability, the ability to hear softer sounds at a particular frequency. The threshold of hearing, the softest sounds that are audible for each frequency, increases as hearing loss progresses. Changes in this threshold can either be a temporary threshold shift (TTS) or a permanent threshold shift (PTS). Often these changes occur in the higher frequencies of 3000 to 6000 Hz, with a notch or significant reduction in hearing ability often around 4000 Hz.

A single exposure to short-duration, extreme loud noise or repeated and prolonged exposure to loud noises are the two most common causes of hearing loss. Examples of the first might be exposure to noise from discharging firearms, while the second might be the cumulative effects of working in a noisy environment such as manufacturing or in loud concert venues. Some antibiotics, drugs, and chemicals can also cause permanent injury.

Hearing damage isn't the only health effect of noise. Workers in noisy workplaces have shown a higher likelihood of heart disease and heart attacks. Numerous other stress-related effects have been documented, including studies that have shown that women in noisy environments tend to gain weight.

47.2 How Loud Is Too Loud? OSHA, NIOSH, EPA, WHO

As in other industries, workers in the sound industry are covered by the occupational noise exposure standard found in the Code of Federal Regulations (29 CFR 1910.95). Occupational Safety and Health (OSHA) regulation requires that workers' exposures not exceed those in Table 47-1.

Table 47-1. Permissible Noise Exposures

Duration per Day, Hours	Sound Level dBA Slow Response
8	90
6	92
4	95
3	97
2	100
1 ½	102
1	105
½	110
¼ or less	115

Noise levels are measured with a sound level meter or dosimeter (a sound level meter worn on the employee) that can automatically determine the average noise level. Often, noise levels are represented in terms of a daily dose. For example, a person who was exposed to an average level of 90 dBA for four hours would have received a 50% dose, or half of her allowable exposure.

Administrative controls—such as the boss saying, "Don't work in noisy areas, or do so for only short times," and/or engineering controls—such as quieter machines—are required to limit exposure. Hearing protection may also be used, although it is not the preferred method. Moreover, the regulation requires that, for employees whose exposure may equal or exceed an 8-hour time-weighted average of 85 dB, the employer shall develop and implement a monitoring program in which employees receive an annual hearing test. The testing must be provided for free to the employee. The employer is also required to provide a selection of hearing protectors and take other measures to protect the worker.

Compliance by employers with the OSHA regulations, as well as enforcement of the regulation, is quite variable, and often it is only in response to requests from employees. It is quite possible that professionals in the field have never had an employer-sponsored hearing test, and are not participating in a hearing conservation program as required.

Unfortunately, OSHA's regulations are among the least protective of any developed nation's hearing protections standards. Scientists and OSHA itself have known for more than a quarter-century that between 20 and 30% of the population exposed to OSHA-permitted noise levels over their lifetime will suffer substantial hearing loss, see Table 47-2. As a result, the National Institute of Occupational Safety and Health (NIOSH), a branch of the Centers for Disease Control and Prevention (CDC), has recommended an 85 dB standard as shown in Table 47-3. Nevertheless, NIOSH recognizes that approximately 10% of the population exposed to the lower recommended level will still develop hearing loss.

Table 47-2. NIOSH's 1997 Study of Estimating Excess Risk of Material Hearing Impairment

Average Exposure Level–dBA	Risk of Hearing Loss Depending on the Definition of Hearing Loss Used
90 (OSHA)	25–32%
85 (NIOSH)	8–14%
80	1–5%

While 25–30% of the population will suffer substantial hearing loss at OSHA permitted levels, everyone would suffer some hearing damage.

Table 47-3 compares the permissible or recommended daily exposure times for noises of various levels. The table is complicated but instructive. The first three columns represent the recommendations of the Environmental Protection Agency (EPA) and World Health Organization (WHO) and starts with the recommendation that the 8-hour average of noise exposure not exceed 75 dBA. The time of exposure is reduced by half for each 3 dBA that is added; a 4-hour exposure is 78 dBA, and a 2-hour exposure is 81 dBA. This is called a 3 dB exchange rate, and is justified on the principle that a 3 dB increase is a doubling of the energy received by the ear, and therefore exposure time ought to be cut in half. The EPA and WHO recommendations can be thought of as safe exposure levels. The NIOSH recommendations in the next three columns represent an increased level of risk of hearing loss and are not protective for approximately 10% of the population. NIOSH uses a 3 dB exchange rate, but the 8-hour exposure is 10 dB higher than EPA—that is, 85 dBA. Finally, the OSHA limits are in the last two columns. OSHA uses a 5 dB exchange rate, which results in much longer exposure times at higher noise levels, and the 8-hour exposure is 90 dBA. Between 20 and 30% of people exposed to OSHA-permitted levels will experience significant hearing loss over a lifetime of expo-

Table 47-3. EPA, WHO, NIOSH, and OSHA Recommended Decibel Standards

	EPA and WHO			NIOSH			OSHA	
dBA	Hours	Min	s	Hours	Min	s	Hours	Min
75	8							
76								
77								
78	4							
79								
80								
81	2							
82								
83								
84	1							
85				8				
86								
87		30						
88				4				
89								
90		15					8	
91				2				
92								
93		7	30					
94				1				
95							4	
96		3	45					
97					30			
98								
99		1	53					
100					15		2	
101								
102			56					
103					7	30		
104								
105			28				1	60
106					3	45		
107								
108			14					
109					1	53		
110							0.5	30
111			7					
112						56		
113								
114			4					
115						28	0.25	15

sure. It is important to note that everyone exposed to the OSHA-permitted levels over their lifetime will experience some hearing loss.

It is important to remember that each of these recommendations assumes that one is accounting for all of the noise exposure for the day. Someone who is working in a noisy environment, then goes home and uses power tools or lawn equipment, is further increasing the risk and exposure.

The U.S. Environmental Protection Agency (EPA) and the World Health Organization (WHO) have recommended a 75 dB limit, as shown in Table 47-3, as a safe exposure with minimal risk of hearing loss. The WHO goes on to recommend that exposure such as at a rock concert be limited to four times per year.

47.3 Indicators of Hearing Damage

There are several indicators of hearing damage. Since the damage is both often slow to manifest itself and progressive, the most important indicators are the ones that can be identified before permanent hearing damage has occurred.

The first and most obvious indicator is exceeding the EPA and WHO safe noise levels. As noise 8 hours, risk of suffering hearing loss also increases.

Exceeding the safe levels by, for example, working at OSHA-permitted noise levels doesn't necessarily mean you will suffer substantial hearing loss; some people will suffer substantial loss, but everyone will suffer some level of hearing damage. The problem is that there is no way to know if you are in the one quarter to one third of the population who will suffer substantial hearing loss at a 90 dBA level or the two thirds to thre quarters of the population who will lose less—at least, not until it is too late and the damage has occurred. Of course, by greatly exceeding OSHA limits, you can be assured that you will have significant hearing loss.

There are two types of temporary hearing damage that are good indicators that permanent damage will occur if exposure continues. The first is tinnitus, a temporary ringing in the ears following a loud or prolonged noise exposure. Work that induces tinnitus is clearly too loud, and steps should immediately be taken to limit exposure in the future.

The second type of temporary damage that is a useful indicator of potential permanent damage is a temporary threshold shift (TTS). Temporary changes in the threshold of hearing, the softest sounds that are audible for each frequency, are a very good indicator that continued noise exposure could lead to permanent hearing loss. Although ways to detect TTS without costly equipment are now being developed, the subjective experience of your hearing sounding different after noise exposure currently provides the best indication of problems.

It is important to remember that the absence of either of these indicators does not mean you will not suffer hearing loss. The presence of either is a good indication that noise exposure is too great.

Regular hearing tests can't detect changes in hearing before they become permanent, but if frequent enough, they can detect changes before they become severe. It is particularly important, therefore, that people exposed to loud noises receive regular hearing tests.

Finally, there are often indicators that serious hearing damage has occurred, such as difficulties understanding people in crowded, noisy situations (loud restaurants, for example), the need to say "What?" frequently, or asking people to repeat themselves. Often it is not the person with the hearing loss, but rather others around him or her, who are the first to recognize these problems due to the slow changes to hearing ability and denial that often accompany them. While it is impossible to reverse hearing damage, hearing loss can be mitigated somewhat by the use of hearing aids, and further damage can be prevented. It is important to remember that just because you have damaged your hearing doesn't mean you can't still make it much worse.

47.4 Protecting Your Hearing

Protecting your hearing is reasonably straightforward: avoid exposure to loud sounds for extended periods of time. This can be accomplished by either turning down the volume or preventing the full energy of the sound from reaching your ears.

There are several strategies for protecting your hearing if you believe or determine that your exposure exceeds safe levels. As Table 47-3 indicates, you can reduce *the noise level* or *reduce the exposure time*, or both.

While reducing exposure time is straightforward it is not always possible, in which case turning down the volume by using quieter equipment, maintaining a greater distance from the noise source, using barriers or noise-absorbing materials, or utilizing hearing protection (either earplugs or over-the-ear muffs, or both) are required.

Typical earplugs or earmuffs are often criticized for changing the sound and hindering communication. Hearing protection in general is far better at reducing noise in the higher frequencies than the lower frequencies, so typical hearing protection significantly changes the sound a wearer is hearing. Consonant sounds in speech occur in the frequencies that are more greatly attenuated by some hearing protectors.

There are, however, a number of hearing protection devices designed to reduce noise levels in all frequencies equally. Often referred to as musician's earplugs, these can come in inexpensive models or custom-molded models. The advantage of a flat or linear attenuation of noise across all frequencies is that the only change to the sound is a reduction in noise level.

47.4.1 Protecting Concert-Goers and Other Listeners

Ears are for listening, and when it comes to music, there are often many ears listening to the music. They too, like music professionals, are at risk of hearing loss. Loud music is exciting; that is the physiology of loud. It gives us a shot of adrenaline. Also, more neurons are firing in our brain and our chest is resonating with the low-frequency sounds.

When humans evolved, the world was much quieter than it is today. Infrequent thunder was about it for loud noise. Hearing evolved to be a very important sense with respect to our survival, working 24/7 to keep us informed about the changing conditions of our environment. Noise wakes us up, because if it didn't wake our forebears up when trouble entered the camp, they might not live long enough to create descendants. Noise is an important warning device—think of a child's crying or screaming. During most of human history, when it was loud, trouble was involved. Physiologically, loud noises give us a shot of adrenaline, gearing us up to either fight or flee. Today, while neither fight nor flight is an appropriate response to loud noise, we still receive that shot of adrenaline. This is the reason for the popularity of loud movie soundtracks, loud exercise gyms, and loud music. It adds excitement and energy to activities. But it is also the reason for the stress-related effects of noise.

There is great incentive to turn it up, especially since the consequences are often not experienced until years later when the extent of hearing damage becomes apparent. People come to concert venues for excitement, not to be bored, and they come willingly; in fact, they pay to inflict whatever damage might be caused. Still, it is not a well-informed decision, and often minors are in the audience. But mostly, it isn't necessary. The desired physiological responses occur at lower noise levels. Moreover, it makes little sense for an industry to degrade the experience of listening to music in the future for whatever marginal gain comes from turning it up a few more decibels now.

Fortunately, even small gestures to turn it down have noticeable impacts. Because every 3 dB decrease halves exposure, small decreases in sound pressure level can vastly increase public safety.

47.4.2 Protecting the Community

Noise can spill over from a venue into the community. The term *noise* has two very different meanings. When discussing hearing loss, noise refers to a sound that is loud enough to risk hearing loss. In a community setting, noise is aural litter. It is audible trash. Noise is to the soundscape as litter is to the landscape. When noise spills over into the community, it is the aural equivalent of throwing McDonald's wrappers onto someone else's property.

When noise reaches the community, often it has lost its higher-frequency content, as that is more easily attenuated by buildings, barriers, and even the atmosphere. What is often left is the bass sound.

Solutions to community noise problems are as numerous as the problems themselves, and usually require the expertise of architectural acousticians. In general, carefully aimed distributed speaker systems are better than large stacks for outdoor venues. Barriers can help, but not in all environmental conditions, and their effectiveness tends to be limited to nearer neighbors. Moreover, barriers need to be well designed, with no gaps.

Indoor walls with higher sound transmission class (STC) ratings are better than ones with lower ratings. STC ratings, however, do not address low-frequency sounds that are most problematic in community noise situations, so professional advice is important when seeking to design better spaces or remedy problems.

Windows and doors are particularly problematic, as even these small openings can negate the effects of very well-soundproofed buildings. They also tend to be the weakest point, even when shut.

Sound absorption is useful for reducing transmission through walls, but in general, decoupling the interior and exterior so that the sound vibrations that hit the interior wall do not cause the exterior wall to vibrate and reradiate the noise is more effective. There are numerous products available to achieve both decoupling and sound absorption.

Often, however, employing these techniques is not an option for the sound engineer. In that case, controlling sound pressure levels and low-frequency levels are the best solution.

47.5 Too Much of a Good Thing

In today's world, noise represents one of the more serious pollutants, Fig. 47-1. Some are the by-product of our society such as lawn mowers, jackhammers, traffic, and public transportation.

ACO Pacific
Noise comes from the Latin "Nausea"
Sounds like the Romans had Noise Problems too...
SM

Figure 47-1. Derivation of noise. Courtesy ACO Pacific.

We deliberately subject ourselves to a Pandora's box of sounds that threaten not only our hearing but our general health. Personal sources like MP3 players, car stereos, or home theaters are sources we can control, yet many remain oblivious to their impact, Fig. 47-2. In the public domain clubs, churches, auditoriums, amphitheaters, and stadiums are part of the myriad of potential threats to hearing health. From a nuisance to a serious health risk, these sources impact attendees, employees, and neighbors alike. As pointed out previously, levels of 105 dBA for 1 hour or less may result in serious and permanent hearing damage. Recent studies have shown other factors such as smoking, drugs of all types, and that overall health appear to accelerate the process.

ACO Pacific
ONE'S MUSIC
is often
Another's Noise
SM

Figure 47-2. Loud sounds from passing cars are often aggravating to passers by. Courtesy ACO Pacific.

High sound levels are just part of the problem. Sound does not stop at the property line. Neighbors and neighborhoods are affected. Numerous studies have shown persistent levels of *noise* affect sleeping patterns, even increase the potential for heart disease. Studies by Johns Hopkins have shown hospital noise impacts patients in the neonatal wards and other patients' recovery time.

Communities all over the world have enacted various forms of noise ordinances. Some address *noise* based on the annoyance factor. Others specify noise limits with sound pressure level (SPL), time of day, and day of the week regulations. The problem, noise (sound), is a transient event. Enforcement and compliance are often very difficult, especially when treated as an annoyance.

47.5.1 A Compliance and Enforcement Tool

There are various tools to monitor noise. One very useful tool is "the SLARM™ by ACO Pacific. The following will use the SLARM™ to explain the importance of noise-monitoring test gear. The SLARM™ tool was developed to meet the needs of the noise abatement market. The SLARM™ performs both compliance and enforcement roles, offering accurate measurement, alarm functions, and very important history.

For the business owner dealing with neighborhood complaints, the SLARM™ provides a positive indication of SPL limits—permitting employees to control the levels or even turn off the sound. The History function offers a positive indication of compliance.

On the enforcement side, no longer does enforcement have to deal with finger-pointing complaints. They now may be addressed hours or days after the event and resolved. There is also the u*niform effect*. Police pull up armed with a sound level meter (SLM) and the volume goes down. Businesses now can demonstrate compliance. Yes—it is an oversimplification— but the concept works. Agreements are worked out. Peace and quiet return to the neighborhood.

47.5.1.1 The SLARMSolution™

The SLARM™ (**S**ound **L**evel **A**larm and **M**onitor) is a package of three basic subsystems in a single standalone device:

1. A sound level meter designed to meet or exceed Type 1 specifications.
2. Programmable threshold detectors providing either SPL or Leq alarm indications.
3. Monitor—a data recorder storing SPL data, and Led values for about 3 weeks on a rolling basis, as well as logging unique Alarm events, scheduled

threshold changes, maintenance events,and calibration information.

The SLARM™ may operate standalone. A PC is not required for normal Alarm operation. The data is maintained using flash and ferro-ram devices.

The SLARM™ provides USB and serial connectivity. It may be connected directly to a PC or via optional accessories directly to an Ethernet or radio link such as Bluetooth™.

PC operation is in conjunction with the included SLARMSoft™ software package.

47.5.1.2 SLARMSoft™ Software Suite

SLARMWatch™. A package with password-protected setup, calibration, downloading, display, and clearing of the SLARM™'s SPL history. The history data may be saved and imported for later review and analysis, Fig. 47-3.

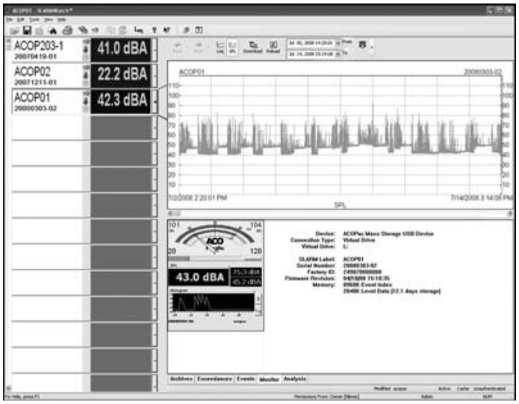
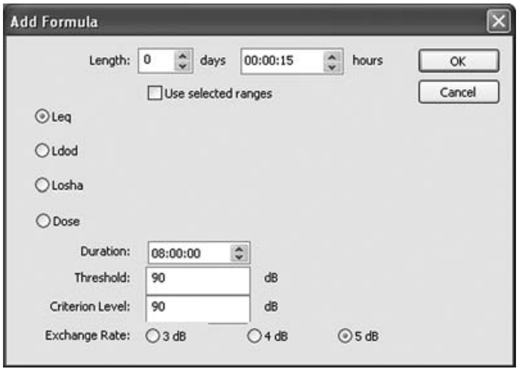

ACOP203-1
41.0 dBA
ACOP02
22.2 dBA
ACOP01
42.3 dBA
ACOP01
43.0 dBA
SPL
Archives
Exceedances
Events
Monitor
Analysis

Figure 47-3. SLARMWATCH™ History and Events and three SLARM™ displays. Courtesy ACO Pacific.

SLARMAnalysis™. Part of SLARMWatch™ provides tools for the advanced user to review the SLARM™ history files. SLARMWatch™ allows saving and storage of this file for later review and analysis. SLARMAnalysis™ provides Leq, Dose and other calculations with user parameters, Fig. 47-4.

SLARMScheduler™. Part of the SLARMWatch™ package, allows 24/7 setting of the Alarm thresholds. This permits time of day and day of the week adjustments to meet the needs of the community, Fig. 47-5.


Add Formula
OK
Cancel

Length: 0 days 00:00:15 hours

☐ Use selected ranges

● Leq

○ Ldod

○ Losha

○ Dose

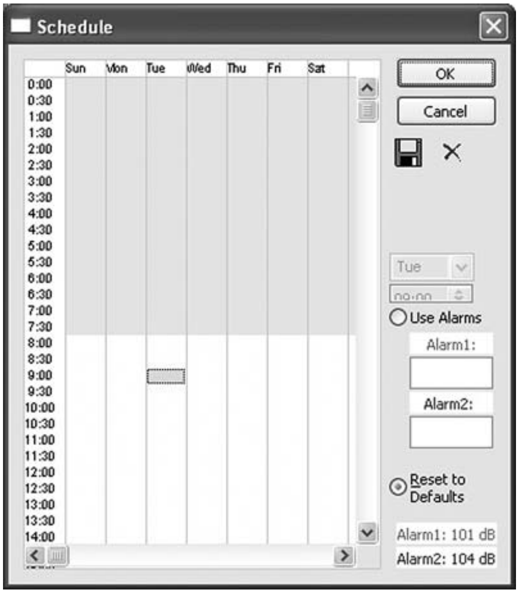
Duration: 08:00:00

Threshold: 90 dB

Criterion Level: 90 dB

Exchange Rate: ○ 3 dB ○ 4 dB ● 5 dB

Figure 47-4. SLARMAnalysis™ Panel Courtesy, ACO Pacific.


Schedule
Sun
Mon
Tue
Wed
Thu
Fri
Sat
0:00
0:30
1:00
1:30
2:00
2:30
3:00
3:30
4:00
4:30
5:00
5:30
6:00
6:30
7:00
7:30
8:00
8:30
9:00
9:30
10:00
10:30
11:00
11:30
12:00
12:30
13:00
13:30
14:00
OK
Cancel
Tue
00:00
Use Alarms
Alarm1:
Alarm2:
Reset to Defaults

Alarm1: 101 dB

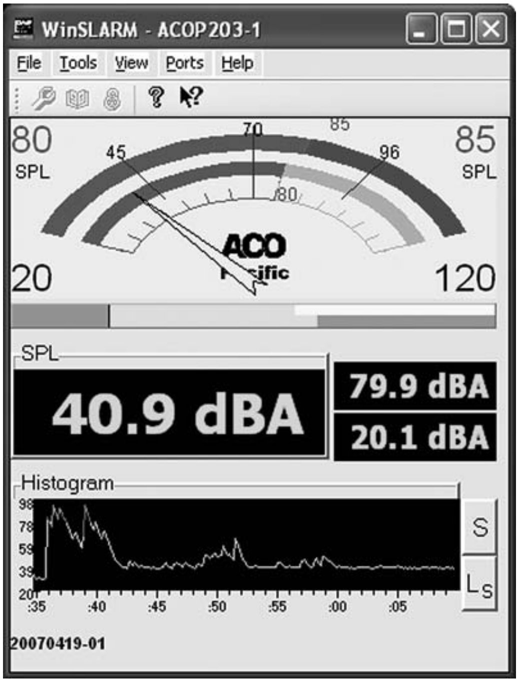
Alarm2: 104 dB

Figure 47-5. SLARMScheduler™ Panel. Thresholds may be individually set for each ALARM over a 24-hour, 7-day period. Courtesy ACO Pacific.

WinSLARM™. A display of SPL, Leqs, Range, and Alarm settings with digital, analog bar graph, and meter displays, as well as a Histogram window that provides a 25 second view of recent SPL on a continuous basis. The WinSlarm™ display may be sized permitting single or multiple SLARM™s to be shown, Fig.47-6.

SLARMAlarm™. Operates independently from SLARMWatch™. The package monitors SLARM™s providing digital display of SPL and Leqs values while also offering SMS, text, and email messaging of Alarm events via an Internet connection from the PC, Fig. 47-7.

SLARMNet™. The SLARM™ and the SLARMSoft™ package allow multiple SLARM™s to be



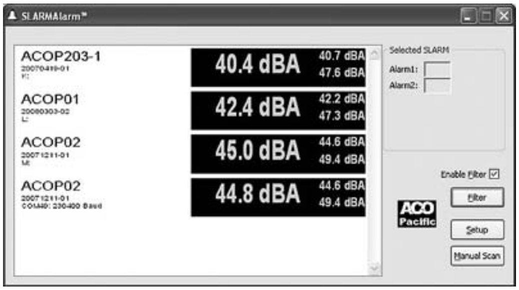
40.9 dBA

79.9 dBA

20.1 dBA

20070419-01

Figure 47-6. WinSLARM™ display provides a real-time look at SPL, Leq Thresholds, and recent events. Courtesy ACO Pacific.



ACOP203-1 40.4 dBA 40.7 dBA 47.6 dBA

ACOP01 42.4 dBA 42.2 dBA 47.3 dBA

ACOP02 45.0 dBA 44.6 dBA 49.4 dBA

ACOP02 44.8 dBA 44.6 dBA 49.4 dBA

Figure 47-7. SLARMAlarm™ display with three SLARM™s. Note: ACOP2 has both USB and Ethernet (via a serial adaptor) connections. Courtesy ACO Pacific.

connected to a network providing real-time data with alarm indications to multiple locations.

47.5.1.3 SLARM™ Operation

The SLARM™ operates in the following manner, Fig. 47-8.

The Microphone and Microphone Preamplifier. The 7052/4052 microphone and preamplifier are supplied with the SLARM™ system. The 7052 is a Type 1.5™ ½ inch free-field measurement microphone featuring a titanium diaphragm. The microphone has a frequency response from <5 Hz to 22 kHz and an output level of 22 mV/Pa (–33 dBV/Pa). The 4052 preamplifier is powered from 12 Vdc supplied by the SLARM™ and has a response <20 Hz to >100 kHz. Together they permit measurements approaching 20 dBA. The MK224 electret capsule is available, offering 8 Hz to 20 kHz response, and 50 mV/Pa (–26 dBV/Pa) performance providing a lower noise floor. The diaphragm is quartz coated nickel.

The Preamplifier (Gain Stage). A low noise gain stage is located after the microphone input. This stage performs two tasks. The first limits the low-frequency input to just under 10 Hz. This reduces low-frequency interference from wind or doors slamming, things we do not hear due to the roll-off of our hearing below 20 Hz. The gain of this stage is controlled by the microcontroller providing two 100 dB measurement ranges 20 to 120 dB and 40 to 140 dBSPL. Most measurements are performed with the 20 to 120 dBSPL ranges. Custom ranges to >170 dBSPL are available as options. The output of the gain stage is supplied to three analog filter stages "A", "C" and "Z" (Linear).

Analog A- and C-Weighted Filters. The gain stage is fed to the C-weighted filter. C-weighted filters have a –3 dB response limit of 31.5 Hz to 8 kHz. C-weighted filters are very useful when resolving issues with low frequencies found in music and industrial applications. The output of the C-weighted filter is connected to both the analog switch providing filter selection and the input of the A-weighted element of the filter system. Sound levels measured with the C-weighted filter are designated as dBC (dBSPL C weighted).

The A-weighted response is commonly found in industrial and community noise ordinances. A weighting rolls off low-frequency sounds. Relative to 1 kHz, the roll-off is –19.4 dB at 100 Hz (a factor of 1:10) and –39.14 at 31.5 Hz (a factor of 1:100). The A response significantly deemphasizes low-frequency sounds. Sound levels measured with the A-weighted filter are designated as dBA (dBSPL A weighted). The output of the A-weighted filter is sent to the analog switch.

Analog Z-Weighting (Linear) Filter. The Z designation basically means the electrical output of the microphone is not weighted. The SLARM™ Z-weighting response is 2 Hz to >100 kHz. The response of the system is essentially defined by the response of the microphone and preamp. Z weighting is useful where measurements of frequency response are desired, or

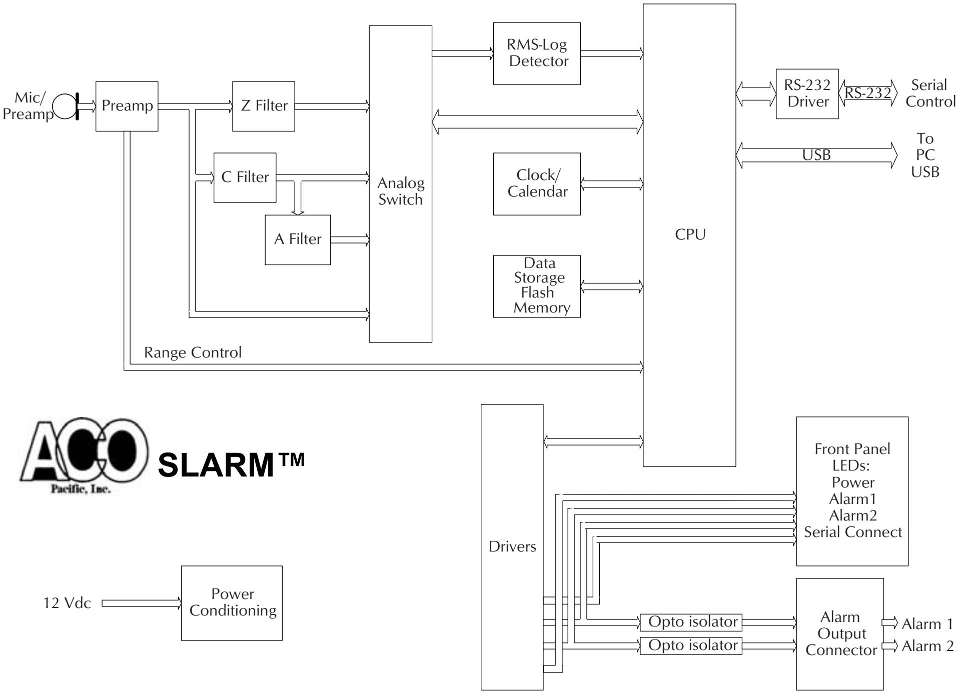

Mic/
Preamp
Preamp
Z Filter
C Filter
A Filter
Analog Switch
RMS-Log Detector
Clock/ Calendar
Data Storage Flash Memory
CPU
RS-232 Driver
RS-232
Serial Control
USB
To PC USB
Range Control
ACO Pacific, Inc.
SLARM™
Drivers
Front Panel LEDs:
Power
Alarm1
Alarm2
Serial Connect
12 Vdc
Power Conditioning
Opto isolator
Opto isolator
Alarm Output Connector
Alarm 1
Alarm 2

Figure 47-8. SLARM™ functional block diagram. Courtesy ACO Pacific.

where low or high frequencies are important. Remember the microphone response determines the response. Sound levels measured with the Z weighted filter are designated as dBZ (dBSPL Z weighted).

Analog Switch. The outputs of the A-, C-, and Z-weighted filters connect to the analog switch. The switch is controlled by the microcontroller. The selection of the desired filter is done at setup using the utilities found in SLARMWatch™.

Selection of the filter as with the other SLARM™ settings is password protected. Permission must be assigned to the user by the administrator before selection is possible. This is essential to minimize the possibility of someone changing measurement profiles that may result in improper ALarm activation or inaccurate measurements.

RMS Detection and LOG Conversion. The output of the analog switch goes to the RMS detection and Logarithmic conversion section of the SLARM™. The RMS detector is a true RMS detector able to handle crest factors of 5–10. This is different from an averaging detector set up provide rms values from sine wave (low crest factor) inputs. The response of the detector exceeds the response limits of the SLARM™.

The output of the RMS detector is fed to the Log (Logarithmic) converter. A logarithmic conversion range of over 100 dB is obtained. The logarithmic output then goes to the A/D section of the microcontroller.

Microcontroller. The microcontroller is the digital heart of the SLARM™. A microcontroller (MCU) does all the internal calculations and system maintenance.

SPL, Leq. The digital data from the internal A/D is converted by the MCU to supply dBSPL, and Leq values for both storage in the on-board flash memory and inclusion in the data stream supplied to the USB and serial ports. These are complex mathematical calculations involving log and anti-log conversation and averaging.

The SPL values are converted to a rolling average. The results are sent to the on-board flash memory that maintains a rolling period of about 2 to 3 weeks.

Leq generation in the SLARM™ involves two independent calculations with two programmable periods. A set of complex calculations generates the two Leq values.

Thresholds and Alarms. The results of the Averaging and Leq calculations are compared by the microcontroller with the Threshold levels stored in the on-board ferro-ram. Threshold levels and types—SPL or Leq—are set using the Settings tools provided in SLARMWatch™. These thresholds are updated by the SLARMScheduler™ routine.

If the programmed threshold limits are exceeded the microcontroller generates an output to an external driver IC. The IC decodes the value supplied by the microcontroller, lighting the correct front panel ALARM LED, and also activating an opto-isolator switch. The opto-switch contacts are phototransistors. The transistor turns on when the opto-isolator LED is activated. The result—a contact closure signaling the outside world of the ALarm.

Real-time Clock. The SLARM™ has an on-board real-time clock. Operating from an internal lithium cell, the real-time clock timestamps all of the recorded history, event logging, and controls the SLARMScheduler™ operation. The Settings panel in SLARMWatch™ allows user synchronization with a PC.

Communicating with the Outside World. SLARM™ may be operated Standalone (without a PC). The SLARM™ provides both USB 2.0 and RS232 serial connections. The USB port is controlled by the microcontroller and provides full access to the SLARM settings, History flash memory, and firmware update capability.

The RS232 is a fully compliant serial port capable of up to 230 k Baud. The serial port may be used to monitor the data stream from the SLARM™. The serial port may also be used to control the SLARM™ settings.

Ethernet and Beyond. Utilizing the wide variety of after-market accessories available, the USB and Serial ports of the SLARM™ may be connected to the Ethernet and Internet. RF links like Bluetooth® and WiFi are also possible. Some accessories will permit the SLARM™ to become an Internet accessory without a PC, permitting remote access from around the world.

The SLARMSoft™ package permits the monitoring of multiple SLARM™s through the SLARMNet™. The SLARMAlarm™ software not only provides a simple digital display of multiple SLARM™s also permits transmission of SMS, text and email of ALarm events. This transmission provides the SLARM™ ID, Time, Type, and Level information in a short message. The world is wired.

History. The on-board flash and ferro-ram memories save measurements, events, settings, user access, and the SLARM™ Label. The SLARM™ updates the flash memory every second. SPL/Leq data storage is on a rolling 2 to 3 week basis. ALARM events, user access, and setting changes are also logged. These maybe downloaded, displayed, and analyzed using features found in SLARMWatch™.

47.5.1.3.1 Applications

SLARM™ applications are virtually unlimited. Day–to–day applications are many. Children's day care centers, hospitals, classrooms, offices, clubs, rehearsal halls, auditoriums, amphitheaters, concert halls, churches, health clubs, and broadcast facilities are among the locations benefitting from sound level monitoring. Industrial and community environments include: machine shops, assembly lines, warehouses, marshaling yards, construction sites and local law enforcement of community noise ordinances.

The following are examples of recent SLARMSolution™.

A Healthy Solution. Located in an older building with a lot of flanking problems, the neighbors of a small women's health club were complaining about the music used with the exercise routines. Negotiations were at a standstill until measurements were made.

Music levels were measured in the health club and a mutually acceptable level established. A SLARM™ (operating standalone—no PC) was installed to monitor the sound system and a custom control accessory developed to the customer's specifications. If the desired SPL limits were exceeded for a specific period of time, the SLARM™ disabled the sound system, requiring a manual reset. The result, a Healthy Solution.

Making a Dam Site Safer. A SLARM™ (operating standalone—no PC) combined with an Outdoor Microphone assembly (ODM) located 300 ft away, monitors the 140+ dBSPL of a Gate Warning Horn. The operator over 100 miles away controls the flood gates of the dam, triggering the horn. The PLC controls the gate operation and monitors power to the horn but not the acoustic output. The SLARMSolution™ monitors the sound level from the horn. The thresholds were set for the normal level and a minimum acceptable level. The minimum level alarm or no alarm signal prompts maintenance action. The SLARM™'s history provides proof of proper operation. Alarm events are time-stamped and logged.

Is It Loud Enough? Tornado, fire, nuclear power plant alarms and sirens as well as many other public safety and industrial warning devices can benefit from monitoring. Using the SLARM™'s standalone operation and the ODM microphone assembly make these remote installations feasible.

A Stinky Problem. A Medivac helicopter on its life-saving mission quickly approaches the hospital helipad and sets down. On the ground, the helicopter engines idle, prepared for a quick response to the next emergency.

The problem: the exhaust fumes from the engines drift upward toward the HVAC vents eight stories above. Specialized carbon filters and engineering staff run to the HVAC controls to turn them off—often forgetting to turn them back on, costing the hospital over $50,000 a year and hundreds of manhours provided limited success.

A standalone SLARM™ with an ODM microphone mounted on the edge of the helipad detects arriving helicopters and turns off the HVAC intakes. As the helicopter departs, the vents are turned back on automatically. The SLARM™ not only provides control of the HVAC but also logs the arrival and departure events for future review, Fig. 47-9.

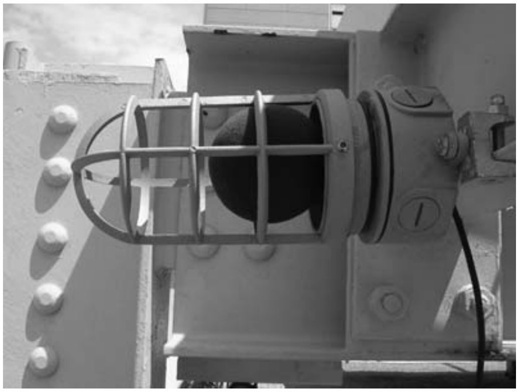


Figure 47-9. ODM microphone assembly mounted on helipad. Courtesy ACO Pacific.

Too Much of a Good Thing Is a Problem. Noise complaints are often the result of *Too Much of a Good Thing*. A nightclub housed on the ground floor of a condo complex faced increased complaints from both condo owners and patrons alike.

The installation of a SLARM™ connected to the DJ's and sound staff's PC allowed them to monitor actual sound levels and alarm them of exceedance. The combination of the SLARM™'s positive indication of compliance and accidence assures maintenance of proper levels.

Protecting the Audience. Community and national regulations often specify noise limits for patrons and employees alike. Faced with the need to assure their audiences' hearing was not damaged by *Too Much of a Good Thing*, a major broadcast company chose the SLARMSolution™.

Two SLARM™s were used to monitor stage and auditorium levels. These units made use of both SPL and Leq Alarm settings. In addition, SLARMAnalysis™ is utilized to extrapolate daily Leq and dose estimates. The installations used the standard SLARM™ mic package and ACO Pacific's 7052PH phantom microphone system. The phantom system utilized the miles of microphone cables running through the complex. This made microphone placement easier. The results were proof of compliance, and the assurance that audience ears were not damaged.

NAMM 2008 – Actual Measurements from the Show Floor. A SLARM™ was installed in a booth at the Winter NAMM 2008 show in Anaheim, CA, The microphone was placed at the back of the booth about 8 ft above the ground away from the booth traffic (people talking).

The following charts utilized SLARMWatch™'s History display capability as well as the SLARMAnalysis™ package. The SLARM™ operated standalone in the booth with the front panel LEDs advising the booth staff of critical noise levels.

The charts show the results of all four days of NAMM and Day 2. Day 2 was extracted from the data using the Zoom feature in SLARMWatch™. The booth was powered down in the evening, thus the Quiet periods shown and the break in the history sequence. The floor traffic quickly picked up at the beginning of the show day.

An 8 hour exposure at these levels has the potential of permanent hearing damage. The booth was located in one of the quieter areas of the NAMM Exhibition floor. Levels on the main show floor were at least 10–15 dB higher than those shown on the graphs.

47.6 Summary

We live in a world of sounds and noise. Some is enjoyable, some annoying, and all potentially harmful to health. Devices like the SLARM™ represent a unique approach to sound control and monitoring and a useful tool for sound and noise pollution control. We hope we have provided insight into how much sound—noise to some—is part of our world to enjoy responsibly, Also so alerting you to the potential harm sound represents.

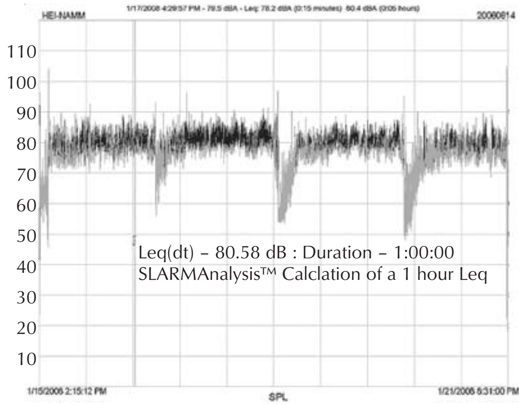


Figure 47-10. This is a dBA (A weighted SPL) for all 4 days of NAMM- The booth power was shut down in the evening and then turned on for the exhibition. The SLARM™ restarted itself each morning and logged automatically during this time. It was not connected to a computer. The black indications are of sound levels exceeding the thresholds set in the SLARM™. Courtesy ACO Pacific.

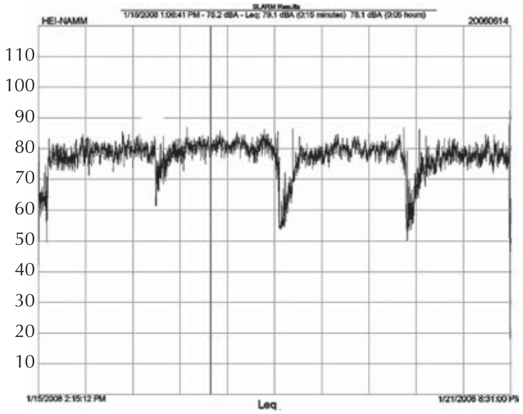


Figure 47-11. All four days 15 s LeqA. Courtesy ACO Pacific.

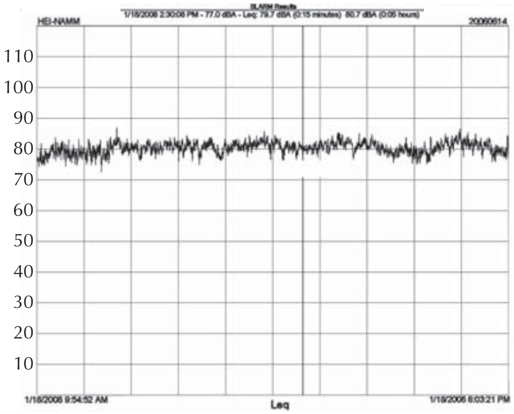


Figure 47-12. Day 2—a typical day. This chart is the Leq (15 s) dBA. This basically represents the running average sound level. Courtesy ACO Pacific.

This page intentionally left blank

Chapter 48

Fundamentals and Units of Measurement

by Glen Ballou

48.1 Units of Measurement 1647
48.1.1 SI System 1647
48.1.2 Fundamental Quantities 1647
48.1.3 Derived Quantities 1648
48.1.4 Definition of the Quantities 1648
48.2 Radio Frequency Spectrum 1654
48.3 Decibel (dB) 1655
48.4 Sound Pressure Level 1657
48.5 Sound System Quantities and Design Formulas 1658
48.6 ISO Numbers 1662
48.7 Greek Alphabet 1663
48.8 Audio Standards 1663
48.9 Audio Frequency Range 1666
48.10 Common Conversion Factors 1666
48.11 Technical Abbreviations 1667
48.12 Audio Frequency Range 1678
48.13 Surface Area and Volume Equations 1678

This page intentionally left blank

48.1 Units of Measurement

Measurements are the method we use to define all things in life. A dimension is any measurable extent such as length, thickness, or weight. A measurement system is any group of related unit names that state the quantity of properties for the items we see, taste, hear, smell, or touch.

A unit of measurement is the size of a quantity in the terms of which that quantity is measured or expressed, for instance, inches, miles, centimeters, and meters.

The laws of physics, which includes sound, are defined through dimensional equations that are defined from their units of measurements of mass, length, and time. For instance,

$$Area = L \times W$$

$$Velocity = \frac{D}{T}$$

where
L is length,
W is width,
D is distance,
T is time.

A physical quantity is specified by a number and a unit, for instance: 16 ft or 5 m.

48.1.1 SI System

The SI system (from the French Système International d'Unités) is the accepted international modernized metric system of measurement. It is used worldwide with the exception of a few countries including the United States of America.

The SI system has the following advantages:

1. Internationally accepted.
2. All values, except time, are decimal multiples or submultiples of the basic unit.
3. It is easy to use.
4. It is easy to teach.
5. It improves international trade and understanding.
6. It is coherent. All derived units are formed by multiplying and dividing other units without introducing any numerical conversion factor except one.
7. It is consistent. Each physical quantity has only one primary unit associated with it.

When using the SI system, exponents or symbol prefixes are commonly used. Table 48-1 is a chart of the accepted name of the number, its exponential form, symbol, and prefix name. (Note because of their size, the numbers from sextillion to centillion have not been shown in numerical form and symbols and prefix names have not been established for these numbers.)

Table 48-1. Multiple and Submultiple Prefixes

Name of Number	Number	Exponential Form	Symbol	Prefix
Centillion		1.0×10^{303}		
Googol		1.0×10^{100}		
Vigintillion		1.0×10^{63}		
Novemdecillion		1.0×10^{60}		
Octodecillion		1.0×10^{57}		
Septendecillion		1.0×10^{54}		
Sexdecillion		1.0×10^{51}		
Quindecillion		1.0×10^{48}		
Quattuordecillion		1.0×10^{45}		
Tredecillion		1.0×10^{42}		
Duodecillion		1.0×10^{39}		
Undecillion		1.0×10^{36}		
Decillion		1.0×10^{33}		
Nonillion		1.0×10^{30}		
Octillion		1.0×10^{27}		
Septillion		1.0×10^{24}	E	Exa-
Sextillion		1.0×10^{21}	P	Peta-
Trillion	1,000,000,000,000	1.0×10^{12}	T	Tera-
Billion	1,000,000,000	1.0×10^{9}	G	Giga-
Million	1,000,000	1.0×10^{6}	M	Mega-
Thousand	1000	1.0×10^{3}	k	Kilo-
Hundred	100	1.0×10^{2}	h	Hecto-
Ten	10	1.0×10^{1}	da	Deka-
Unit	1	1.0×10^{0}	—	—
Tenth	0.10	1.0×10^{-1}	d	Deci-
Hundredth	0.01	1.0×10^{-2}	c	Centi-
Thousandth	0.00	1.0×10^{-3}	m	Milli-
Millionth	0.000 001	1.0×10^{-6}	µ	Micro-
Billionth	0.000 000 001	1.0×10^{-9}	n	Nano-
Trillionth	0.000 000 000 001	1.0×10^{-12}	p	Pico-
Quadrillionth	0.000 000 000 000 001	1.0×10^{-15}	f	Femto-

48.1.2 Fundamental Quantities

There are seven fundamental quantities in physics: length, mass, time, intensity of electric current, temperature, luminous intensity, and molecular substance. Two supplementary quantities are plane angle and solid angle.

48.1.3 Derived Quantities

Derived quantities are those defined in terms of the seven fundamental quantities, for instance, speed = length/time. There are sixteen derived quantities with names of their own: energy (work, quantity of heat), force, pressure, power, electric charge, electric potential difference (voltage), electric resistance, electric conductance, electric capacitance, electric inductance, frequency, magnetic flux, magnetic flux density, luminous flux, illuminance, and customary temperature. Following are thirteen additional derived quantities that carry the units of the original units that are combined. They are area, volume, density, velocity, acceleration, angular velocity, angular acceleration, kinematic viscosity, dynamic viscosity, electric field strength, magnetomotive force, magnetic field strength, and luminance.

48.1.4 Definition of the Quantities

The quantities will be defined in SI units, and their U.S. customary unit equivalent values will also be given.

Length (L). *Length* is the measure of how long something is from end to end. The meter (abbreviated m) is the SI unit of length. (Note: in the United States the spelling “meter” is retained, while most other countries use the spelling “metre.”) The meter is the 1 650 763.73 wavelengths, in vacuum, of the radiation corresponding to the unperturbed transition between energy level $2P_{10}$ and $5D_5$ of the krypton-86 atom. The result is an orange-red line with a wavelength of 6057.802×10^{-10} meters. The meter is equivalent to 39.370 079 inches.

Mass (M). *Mass* is the measure of the inertia of a particle. The mass of a body is defined by the equation

$$M = \left(\frac{A_s}{a}\right)M_s \qquad (48\text{-}1)$$

where,

A_s is the acceleration of the standard mass M_s,

a is the acceleration of the unknown mass, M, when the two bodies interact.

The kilogram (kg) is the unit of mass. This is the only base or derived unit in the SI system that contains a prefix. Multiples are formed by attaching prefixes to the word gram. Small masses may be described in grams (g) or milligrams (mg) and large masses in megagrams. Note the term *tonnes* is sometimes used for the metric ton or megagram, but this term is not recommended.

The present international definition of the kilogram is the mass of a special cylinder of platinum iridium alloy maintained at the International Bureau of Weights and Measures, Sevres, France. One kilogram is equal to 2.204 622 6 avoirdupois pounds (lb). A liter of pure water at standard temperature and pressure has a mass of 1 kg ± one part in 10^4.

Mass of a body is often revealed by its weight, which the gravitational attraction of the earth gives to that body.

If a mass is weighed on the moon, its mass would be the same as on earth, but its weight would be less due to the small amount of gravity.

$$M = \frac{W}{g} \qquad (48\text{-}2)$$

where,

W is the weight,

g is the acceleration due to gravity.

Time (t). *Time* is the period between two events or the point or period during which something exists, happens, etc.

The second (s) is the unit of time. Time is the one dimension that does not have powers of ten multipliers in the SI system. Short periods of time can be described in milliseconds (ms) and microseconds (μs). Longer periods of time are expressed in minutes (1 min = 60 s) and hours (1 h = 3600 s). Still longer periods of time are the day, week, month, and year. The present international definition of the second is the time duration of 9, 192, 631, 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the atom of caesium 133. It is also defined as 1/86, 400 of the mean solar day.

Current (I). *Current* is the rate of flow of electrons. The ampere (A) is the unit of measure for current. Small currents are measured in milliamperes (mA) and microamperes (μA), and large currents are in kiloamperes (kA). The international definition of the ampere is the constant current that, if maintained in two straight parallel conductors of infinite length and negligible cross-sectional area and placed exactly 1 m apart in a vacuum, will produce between them a force of 2×10^{-7} N/m^2 of length.

A simple definition of one ampere of current is the intensity of current flow through a 1 ohm resistance under a pressure of 1 volt of potential difference.

Temperature (T). *Temperature* is the degree of hotness or coldness of anything. The kelvin (K) is the unit of temperature. The kelvin is 1/273.16 of the thermodynamic temperature of the triple point of pure water. Note: the term degree (°) is not used with the term kelvin as it is with other temperature scales.

Ordinary temperature measurements are made with the celsius scale on which water freezes at 0°C and boils at 100°C. A change of 1°C is equal to a change of 1 kelvin, therefore 0°C = 273.15 K: 0°C = 32°F.

Luminous Intensity (I_L). *Luminous intensity* is the luminous flux emitted per unit solid angle by a point source in a given direction. The candela (cd) is the unit of luminous intensity. One candela will produce a luminous flux of 1 lumen within a solid angle of 1 steradian.

The international definition of the candela is the luminous intensity, perpendicular to the surface, of 1⁄600 000 m^2 of a black body at the temperature of freezing platinum under a pressure of 101 325 N/m^2 (pascals).

Molecular Substance (*n*). *Molecular substance* is the amount of substance of a system that contains as many elementary entities as there are atoms in 0.012 kg of carbon 12.

The mole is the unit of molecular substance. One mole of any substance is the gram molecular weight of the material. For example, 1 mole of water (H_20) weighs 18.016 g.

$$\begin{aligned} H_2 &= 2 \text{ atoms} \times 1.008 \text{ atomic weight} \\ O &= 2 \text{ atoms} \times 16 \text{ atomic weight} \\ H_2O &= 18.016\ g \end{aligned}$$

Plane Angle (α). The *plane angle* is formed between two straight lines or surfaces that meet. The radian (rad) is the unit of plane angles. One radian is the angle formed between two radii of a circle and subtended by an arc whose length is equal to the radius. There are 2π radians in 360°.

Ordinary measurements are still made in degrees. The degree can be divided into minutes and seconds or into tenths and hundredths of a degree. For small angles, the latter is most useful.

$$\begin{aligned} \text{One degree of arc } (1°) &= \frac{\pi}{180}\text{Rad} \\ 1\text{Rad} &= 57.2956° \end{aligned} \qquad (48\text{-}3)$$

Solid Angle (*A*). A *solid angle* subtends three dimensions. The solid angle is measured by the area, subtended (by projection) on a sphere of unit radius by the ratio of the area *A*, intercepted on a sphere of radius *r* to the square of the radius (A/r^2).

The steradian (*sr*) is the unit of solid angle. The steradian is the solid angle at the center of a sphere that subtends an area on the spherical surface, which is equal to that of a square whose sides are equal to the radius of the sphere.

Energy (E). *Energy* is the property of a system that is a measure of its ability to do work. There are two main forms of energy—potential energy and kinetic energy.

1. Potential energy (*U*) is the energy possessed by a body or system by virtue of position and is equal to the work done in changing the system from some standard configuration to its present state. Potential energy is calculated with the equation

   $$U = Mgh \qquad (48\text{-}4)$$

   where,
   M is the mass,
   g is the acceleration due to gravity.
   h is the height.

   For example, a mass *M* placed at a height h above a datum level in a gravitational field with an acceleration of free fall (*g*), has a potential energy given by $U = mgh$. This potential energy is converted into kinetic energy when the body falls between the levels.
2. Kinetic energy (*T*) is the energy possessed by virtue of motion and is equal to the work that would be required to bring the body to rest. A body undergoing translational motion with velocity, *v*, has a kinetic energy given by

   $$T = 0.5Mv^2 \qquad (48\text{-}5)$$

   where,
   M is the mass of the body,
   v is the velocity of the body.

   For a body undergoing rotational motion

   $$T = 0.51I\omega^2 \qquad (48\text{-}6)$$

   where,
   I is the moment of inertia of the body about its axis of rotation,
   ω is the angular velocity.

The joule (J) is the unit of energy. The mechanical definition is the work done when the force of 1 newton is applied for a distance of 1 m in the direction of its application, or 1 Nm. The electrical unit of energy is the kilowatt-hour (kWh), which is equal to 3.6×10^6 J.

In physics, the unit of energy is the electron volt (eV), which is equal to $(1.602\ 10 \pm 0.000\ 07) \times 10^{-19}$ J.

Force (*F*). *Force* is any action that changes, or tends to change, a body's state of rest or uniform motion in a straight line.

The newton (*N*) is the unit of force and is that force which, when applied to a body having a mass of 1 kg, gives it an acceleration of 1 m/s^2. One newton equals 1 J/m, 1 kg(m)/s^2, 10^5 dynes, and 0.224 809 lb force.

Pressure. *Pressure* is the force (in a fluid) exerted per unit area on an infinitesimal plane situated at the point. In a fluid at rest, the pressure at any point is the same in all directions. A fluid is any material substance which in static equilibrium cannot exert tangential force across a surface but can exert only pressure. Liquids and gases are fluids.

The pascal (Pa) is the unit of pressure. The pascal is equal to the newton per square meter (N/m^2).

$$\begin{aligned}1Pa &= 10^{-6}\text{bars}\\ &= 1.45038 \times 10^{-4}\text{lb/in}^2\end{aligned} \tag{48-7}$$

Power (*W*). *Power* is the rate at which energy is expended or work is done. The watt (W) is the unit of power and is the power that generates energy at the rate of 1 J/s.

$$\begin{aligned}1\text{W} &= 1\text{J/s}\\ &= 3.141442\ \text{BTU/h}\\ &= 44.2537\ \text{ft-lb/min}\\ &= 0.00134102\ \text{hp}\end{aligned} \tag{48-8}$$

Electric Charge (*Q*). *Electric charge* is the quantity of electricity or electrons that flows past a point in a period of time. The coulomb (C) is the unit of electric charge and is the quantity of electricity moved in 1 second by a current of 1 ampere. The coulomb is also defined as 6.24196×10^{18} electronic charges.

Electric Potential Difference (*V*). Often called electromotive force (*emf*) and voltage (*V*), *electric potential difference* is the line integral of the electric field strength between two points. The volt (V) is the unit of electric potential. The volt is the potential difference that will cause a current flow of 1 A between two points in a circuit when the power dissipated between those two points is 1 W.

A simpler definition would be to say a potential difference of 1 V will drive a current of 1 A through a resistance of 1 Ω.

$$\begin{aligned}Volt\ (V) &= \frac{W}{A}\\ &= \frac{J}{A(s)}\\ &= \frac{\text{kg(m}^2)}{s^3A}\\ &= \text{A}\Omega\end{aligned} \tag{48-9}$$

Electric Resistance (*R*). *Electric resistance* is the property of conductors that, depending in their dimensions, material, and temperature, determines the current produced by a given difference of potential. It is also that property of a substance that impedes current and results in the dissipation of power in the form of heat.

The ohm (Ω) is the unit of resistance and is the resistance that will limit the current flow to 1 A when a potential difference of 1 V is applied to it.

$$\begin{aligned}R &= \frac{V}{A}\\ &= \frac{kg(m^2)}{s^3A^3}\end{aligned} \tag{48-10}$$

Electric Conductance (*G*). *Electric conductance* is the reciprocal of resistance. The siemens (S) is the unit of electric conductance. A passive device that has a conductance of 1 S will allow a current flow of 1 A when 1 V potential is applied to it.

$$\begin{aligned}S &= \frac{1}{\Omega}\\ &= \frac{A}{V}\end{aligned} \tag{48-11}$$

Electric Capacitance (*C*). *Electric capacitance* is the property of an isolated conductor or set of conductors and insulators to store electric charge. The farad (F) is the unit of electric capacitance and is defined as the capacitance that exhibits a potential difference of 1 V when it holds a charge of 1 C.

$$\begin{aligned}F &= \frac{C}{V}\\ &= \frac{AS}{V}\end{aligned} \tag{48-12}$$

where,

C is the electric charge in coulombs,

V is the electric potential difference in volts,

A is the current in amperes,

S is the conductance in siemens.

Electric Inductance (L). *Electric inductance* is the property that opposes any change in the existing current. Inductance is only present when the current is changing. The henry (H) is the unit of inductance and is the inductance of a circuit in which an electromotive force of 1 V is developed by a current change of 1 A/s.

$$H = \frac{Vs}{A} \tag{48-13}$$

Frequency (f). *Frequency* is the number of recurrences of a periodic phenomenon in a unit of time. The hertz (Hz) is the unit of frequency and is equal to one cycle per second, 1 Hz = 1 cps. Frequency is often measured in hertz (Hz), kilohertz (kHz), and megahertz (MHz).

Sound Intensity (W/m^2). *Sound intensity* is the rate of flow of sound energy through a unit area normal to the direction of flow. For a sinusoidally varying sound wave the intensity I is related to the sound pressure p and the density ß of the medium by

$$I = \frac{p^2}{\beta c} \tag{48-14}$$

where,
c is the velocity of sound.

The watt per square meter (W/m^2) is the unit of sound intensity.

Magnetic Flux (ϕ). *Magnetic flux* is a measure of the total size of a magnetic field. The weber (Wb) is the unit of magnetic flux, and is the amount of flux that produces an electromotive force of 1 V in a one-turn conductor as it reduces uniformly to zero in 1 s.

$$\begin{aligned} \mathrm{Wb} &= \mathrm{W(s)} \\ &= 10^8 \text{lines of flux} \\ &= \frac{\mathrm{kg(m^2)}}{\mathrm{s^2A}}. \end{aligned} \tag{48-15}$$

Magnetic Flux Density (ß). The *magnetic flux density* is the flux passing through the unit area of a magnetic field in the direction at right angles to the magnetic force. The vector product of the magnetic flux density and the current in a conductor gives the force per unit length of the conductor.

The tesla (T) is the unit of magnetic flux density and is defined as a density of 1 Wb/m^2.

$$\begin{aligned} \mathrm{T} &= \frac{\mathrm{Wb}}{\mathrm{m^2}} \\ &= \frac{\mathrm{V(s)}}{\mathrm{m^2}} \\ &= \frac{\mathrm{kg}}{\mathrm{s^2A}}. \end{aligned} \tag{48-16}$$

Luminous Flux (Φ_v). *Luminous flux* is the rate of flow of radiant energy as evaluated by the luminous sensation that it produces. The lumen (lm) is the unit of luminous flux, which is the amount of luminous flux emitted by a uniform point source whose intensity is 1 steradian.

$$\begin{aligned} \mathrm{lm} &= cd\left(\frac{sr}{m^2}\right) \\ &= 0.0795774 \text{ candlepower} \end{aligned} \tag{48-17}$$

where,
cd is the luminous intensity in candelas,
sr is the solid angle in steradians.

Luminous flux Density (E_v). The *luminous flux density* is the luminous flux incident on a given surface per unit area. It is sometimes called illumination or intensity of illumination. At any point on a surface, the illumination is given by

$$E_v = \frac{\mathrm{d}\Phi_v}{\mathrm{dA}} \tag{48-18}$$

The lux (lx) is the unit of luminous flux density, which is the density of radiant flux of lm/m^2,

$$\begin{aligned} lx &= \frac{lm}{m^2} \\ &= cd\frac{sr}{m^2} \\ &= 0.0929030\,\mathrm{fc} \end{aligned} \tag{48-19}$$

Displacement. *Displacement* is a change in position or the distance moved by a given particle of a system from its position of rest, when acted on by a disturbing force.

Speed/Velocity. *Speed* is the rate of increase of distance traveling by a body. Average speed is found by the equation

$$S = \frac{l}{t} \tag{48-20}$$

where,

S is the speed,
l is the length or distance,
t is the time to travel.

Speed is a scalar quantity as it is not referenced to direction. Instantaneous speed = dl/dt. *Velocity* is the rate of increase of distance traversed by a body in a particular direction.

Velocity is a vector quantity as both speed and direction are indicated. The l/t can often be the same for the velocity and speed of an object, however, when speed is given, the direction of movement is not known. If a body describes a circular path and each successive equal distances along the path is described in equal times, the speed would be constant but the velocity would constantly change due to the change in direction.

Weight. *Weight* is the force exerted on a mass by the gravitational pull of the planet, star, moon, etc., that the mass is near. The weight experienced on earth is due to the earth's gravitational pull, which is 9.806 65 m/s^2, and causes an object to accelerate toward earth at a rate of 9.806 65 m/s^2 or 32 ft/s^2.

The weight of a mass M is $M(g)$. If M is in kg and g in m/s^2, the weight would be in newtons (N). Weight in the U.S. system is in pounds (lb).

Acceleration. *Acceleration* is the rate of change in velocity or the rate of increase or decrease in velocity with time. Acceleration is expressed in meters per second squared (m/s^2), or ft/s^2 in the U.S. system.

Amplitude. *Amplitude* is the magnitude of variation in a changing quantity from its zero value. Amplitude should always be modified with adjectives such as peak, rms, maximum, instantaneous, etc.

Wavelength (*M*). In a periodic wave, the distance between two points of the corresponding phase of two consecutive cycles is the *wavelength*. Wavelength is related to the velocity of propagation (c) and frequency (f) by the equation

$$\lambda = \frac{c}{f} \tag{48-21}$$

The wavelength of a wave traveling in air at sea level and standard temperature and pressure (STP) is

$$\lambda = \frac{331.4\ \text{m/s}}{f} \tag{48-22}$$

or

$$\lambda = \frac{1087.42\ \text{ft/s}}{f} \tag{48-23}$$

For instance, the length of a 1000 Hz wave would be 0.33 m, or 1.09 ft.

Phase. *Phase* is the fraction of the whole period that has elapsed, measured from a fixed datum. A sinusoidal quantity may be expressed as a rotating vector OA. When rotated a full 360 degrees, it represents a sine wave. At any position around the circle, OX is equal in length but said to be X degrees out of phase with OA.

It may also be stated that the phase difference between OA and OX is α. When particles in periodic motion due to the passage of a wave are moving in the same direction with the same relative displacement, they are said to be in phase. Particles in a wave front are in the same phase of vibration when the distance between consecutive wave fronts is equal to the wavelength. The phase difference of two particles at distances X_1 and X_2 is

$$\alpha = \frac{2\pi(X_2 - X_1)}{\lambda} \tag{48-24}$$

Periodic waves, having the same frequency and waveform, are said to be in phase if they reach corresponding amplitudes simultaneously.

Phase Angle. The angle between two vectors representing two periodic functions that have the same frequency is the *phase angle*. Phase angle can also be considered the difference, in degrees, between corresponding stages of the progress of two cycle operations.

Phase Difference (ϕ). *Phase difference* is the difference in electrical degrees or time, between two waves having the same frequency and referenced to the same point in time.

Phase Shift. Any change that occurs in the phase of one quantity or in the phase difference between two or more quantities is the *phase shift*.

Phase Velocity. The *phase velocity* is when a point of constant phase is propagated in a progressive sinusoidal wave.

Temperature. *Temperature* is the measure of the amount of coldness or hotness. While kelvin is the SI standard, temperature is commonly referenced to °C (degrees Celsius) or °F (degrees Fahrenheit).

The lower fixed point (the ice point) is the temperature of a mixture of pure ice and water exposed to the air at standard atmospheric pressure.

The upper fixed point (the steam point) is the temperature of steam from pure water boiling at standard atmospheric pressure.

In the Celsius scale, named after Anders Celsius (1701–1744) and originally called Centigrade, the fixed points are 0°C and 100°C. This scale is used in the SI system.

The Fahrenheit scale, named after Gabriel Daniel Fahrenheit in 1714, has the fixed points at 32°F and 212°F.

To interchange between °C and °F, use the following equations.

$$°C = (°F - 32°) \times \frac{5}{9}$$

$$°F = \left(°C \times \frac{9}{5}\right) + 32° \qquad (48\text{-}25)$$

The absolute temperature scale operates from absolute zero of temperature. Absolute zero is the point where a body cannot be further cooled because all the available thermal energy is extracted.

Absolute zero is 0 kelvin (0 K) or 0° Rankine (0°R). The Kelvin scale, named after Lord Kelvin (1850), is the standard in the SI system and is related to °C.

$$0°C = 273.15K$$

The Rankine scale is related to the Fahrenheit system.

$$32°F = 459.67°R$$

The velocity of sound is affected by temperature. As the temperature increases, the velocity increases. The approximate formula is

$$C = 331.4 \text{ m/s} + 0.607\,T \quad \text{SI units} \qquad (48\text{-}26)$$

where,
T is the temperature in °C.

$$C = 1052 \text{ ft/s} \times 1.106\,T \quad \text{U.S. units} \qquad (48\text{-}27)$$

where,
T is the temperature in °F.

Another simpler equation to determine the velocity of sound is

$$C = 49.00\sqrt{459.69° + °F} \qquad (48\text{-}28)$$

Things that can affect the speed of sound are the sound wave going through a temperature barrier or going through a stream of air such as from an air conditioner. In either case, the wave is deflected the same way that light is refracted in glass.

Pressure and altitude do not affect the speed of sound because at sea level the molecules bombard each other, slowing down their speed. At upper altitudes they are farther apart so they do not bombard each other as often so they reach their destination at the same time.

Thevenin's Theorem. *Thevenin's Theorem* is a method used for reducing complicated networks to a simple circuit consisting of a voltage source and a series impedance. The theorem is applicable to both ac and dc circuits under steady-state conditions.

The theorem states: the current in a terminating impedance connected to any network is the same as if the network were replaced by a generator with a voltage equal to the open-circuit voltage of the network, and whose impedance is the impedance seen by the termination looking back into the network. All generators in the network are replaced with impedance equal to the internal impedances of the generators.

Kirchhoff's Laws. The laws of Kirchhoff can be used for both dc and ac circuits. When used in ac analysis, phase must also be taken into consideration.

Kirchhoff's Voltage Law (KVL). *Kirchhoff's voltage law* states that the sum of the branch voltages for any closed loop is zero at any time. Stated another way, for any closed loop, the sum of the voltage drops equal the sum of the voltage rises at any time.

In the laws of Kirchhoff, individual electric circuit elements are connected according to some wiring plan or schematic. In any closed loop, the voltage drops must be equal to the voltage rises. For example, in the dc circuit of Fig. 48-1, V_1 is the voltage source or rise such as a battery and V_2, V_3, V_4, and V_5 are voltage drops (possibly across resistors) so

$$V_1 = V_2 + V_3 + V_4 + V_5 \qquad (48\text{-}29)$$

or,

$$V_1 - V_2 - V_3 - V_4 - V_5 = 0 \qquad (48\text{-}30)$$

In an ac circuit, phase must be taken into consideration, therefore, the voltage would be

$$V_1 e^{j\omega t} - V_2 e^{j\omega t} - V_3 e^{j\omega t} - V_4 e^{j\omega t} - V_5 e^{j\omega t} = 0 \qquad (48\text{-}31)$$

where,

e^{jwt} is $\cos At + j\sin At$ or Euler's identity.

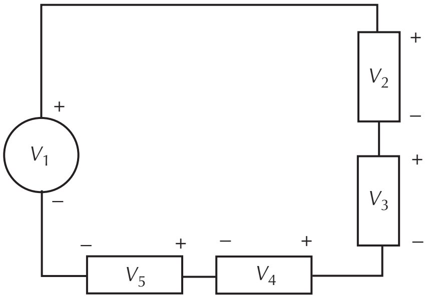


Figure 48-1. Kirchhoff's voltage law.

Kirchhoff's Current Law (KCL). *Kirchhoff's current law* states that the sum of the branch currents leaving any node must equal the sum of the branch currents entering that node at any time.

Stated another way, the sum of all branch currents incident at any node is zero.

In Fig. 48-2 the connection on node current in a dc circuit is equal to 0 and is equal to the sum of currents I_1, I_2, I_3, I_4, and I_5 or

$$I_1 = I_2 + I_3 + I_4 + I_5 \tag{48-32}$$

or

$$I_1 - I_2 - I_3 - I_4 - I_5 = 0 \tag{48-33}$$

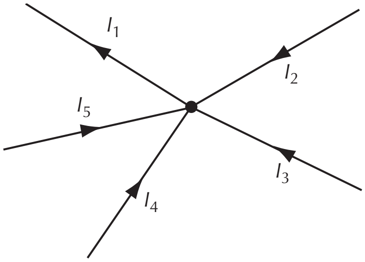


Figure 48-2. Kirchhoff's current law.

The current throughout the circuit is also a function of the current from the power source (V_1) and the current through all of the branch circuits.

In an ac circuit, phase must be taken into consideration, therefore, the current would be

$$I_1 e^{j\omega t} - I_2 e^{j\omega t} - I_3 e^{j\omega t} - I_4 e^{j\omega t} - I_5 e^{j\omega t} = 0 \tag{48-34}$$

where,

e^{jwt} is $\cos At + j\sin At$ or Euler's identity.

Ohm's Law. *Ohm's Law* states that the ratio of applied voltage to the resultant current is a constant at every instant and that this ratio is defined to be the resistance.

If the voltage is expressed in volts and the current in amperes, the resistance is expressed in ohms. In equation form it is

$$R = \frac{V}{I} \tag{48-35}$$

or

$$R = \frac{e}{i} \tag{48-36}$$

where,

e and i are instantaneous voltage and current,

V and I are constant voltage and current,

R is the resistance.

Through the use of Ohm's Law, the relationship between voltage, current, resistance or impedance, and power can be calculated.

Power is the rate of doing work and can be expressed in terms of potential difference between two points (voltage) and the rate of flow required to transform the potential energy from one point to the other (current). If the voltage is in volts or J/C and the current is in amperes or C/s, the product is joules per second or watts:

$$P = \mathrm{V}I \tag{48-37}$$

or

$$\frac{J}{s} = \frac{J}{C}\left(\frac{C}{s}\right) \tag{48-38}$$

where,

J is energy in joules,

C is electric charge in coulombs.

Fig. 48-3 is a wheel chart that relates current, voltage, resistance or impedance, and power. The power factor (PF) is cos I where I is the phase angle between e and i. A power factor is required in ac circuits.

48.2 Radio Frequency Spectrum

The *radio frequency spectrum* of 30 Hz–3,000,000 MHz (3×10^{12} Hz) is divided into the various bands shown in Table 48-2.

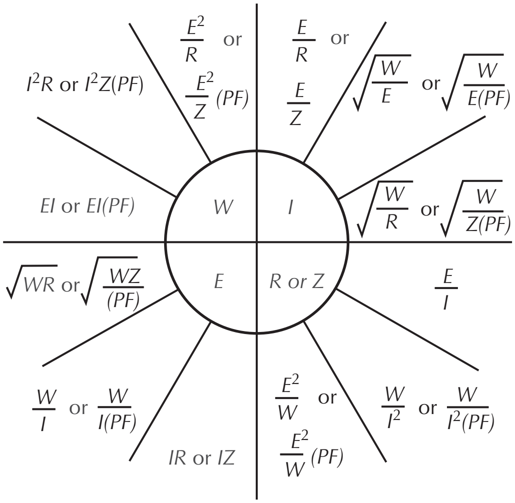

$\frac{E^2}{R}$ or $\frac{E^2}{Z}(PF)$
$\frac{E}{R}$ or $\frac{E}{Z}$
I^2R or $I^2Z(PF)$
$\sqrt{\frac{W}{E}}$ or $\sqrt{\frac{W}{E(PF)}}$
EI or $EI(PF)$
W
I
$\sqrt{\frac{W}{R}}$ or $\sqrt{\frac{W}{Z(PF)}}$
$\sqrt{WR}$ or $\sqrt{\frac{WZ}{(PF)}}$
E
R or Z
$\frac{E}{I}$
$\frac{W}{I}$ or $\frac{W}{I(PF)}$
IR or IZ
$\frac{E^2}{W}$ or $\frac{E^2}{W}(PF)$
$\frac{W}{I^2}$ or $\frac{W}{I^2(PF)}$

Figure 48-3. Power, voltage, current wheel.

Table 48-2. Frequency Classification

Frequency	Band No.	Classification	Abbreviation
30–300 Hz	2	extremely low frequencies	ELF
300–3000 Hz	3	voice frequencies	VF
3–30 kHz	4	very low frequencies	VLF
30–300 kHz	5	low frequencies	LF
300–3000 kHz	6	medium frequencies	MF
3–30 MHz	7	high frequencies	HF
30–300 MHz	8	very high frequencies	VHF
30–3000 MHz	9	ultrahigh frequencies	UHF
3–30 GHz	10	super-high frequencies	SHF
30–300 GHz	11	extremely high frequencies	EHF
300–3 THz	12	–	–

48.3 Decibel (dB)

Decibels are a logarithmic ratio of two numbers. The decibel is derived from two power levels and is also used to show voltage ratios indirectly (by relating voltage to power). The equations or decibels are

$$Power\ dB = 10\log\frac{P_1}{P_2} \tag{48-39}$$

$$Voltage\ dB_v = 20\log\frac{E_1}{E_2} \tag{48-40}$$

Fig. 48-4 shows the relationship between the power, decibels, and voltage. In the illustration, "dBm" is the decibels referenced to 1 mW.

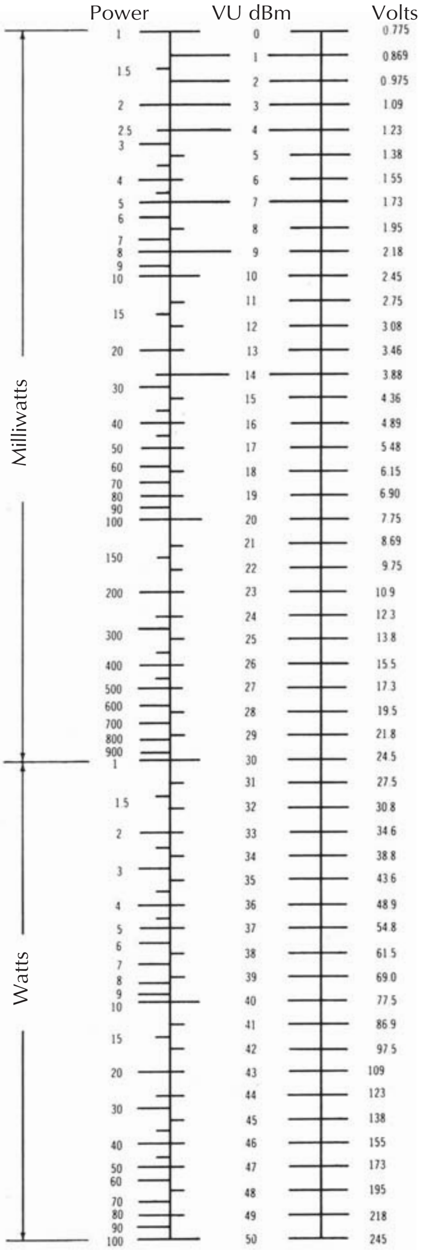


Figure 48-4. Relationship between power, dBm, and voltage.

Table 48-3 shows the relationship between decibel, current, voltage, and power ratios.

Volume unit (VU) meters measure decibels that are related to a 600 Ω impedance, O VU is actually +4 dBm (see Chapter 26). When measuring decibels referenced to 1 mW at any other impedance than 600 Ω, use

Table 48-3. Relationships between Decibel, Current, Voltage, and Power Ratios

dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power
0.0	**1.0000**	**1.000**	**0.0**	**5.0**	**0.5623**	**1.778**	**0.50**	**10.0**	**0.3162**	**3.162**	**5.00**	**15.0**	**0.1778**	**5.623**	**0.50**
0.1	0.9886	1.012	0.05	0.1	0.5559	1.799	0.55	0.1	0.3126	3.199	0.05	0.1	0.1758	5.689	0.55
0.2	0.9772	1.023	0.10	0.2	0.5495	1.820	0.60	0.2	0.3090	3.236	0.10	0.2	0.1738	5.754	0.60
0.3	0.9661	1.035	0.15	0.3	0.5433	1.841	0.65	0.3	0.3055	3.273	0.15	0.3	0.1718	5.821	0.65
0.4	0.9550	1.047	0.20	0.4	0.5370	1.862	0.70	0.4	0.3020	3.311	0.20	0.4	0.1698	5.888	0.70
0.5	0.9441	1.059	0.25	0.5	0.5309	1.884	0.75	0.5	0.2985	3.350	0.25	0.5	0.1679	5.957	0.75
0.6	0.9333	1.072	0.30	0.6	0.5248	1.905	0.80	0.6	0.2951	3.388	0.30	0.6	0.1660	6.026	0.80
0.7	0.9226	1.084	0.35	0.7	0.5188	1.928	0.85	0.7	0.2917	3.428	0.35	0.7	0.1641	6.095	0.85
0.8	0.9120	1.096	0.40	0.8	0.5129	1.950	0.90	0.8	0.2884	3.467	0.40	0.8	0.1622	6.166	0.90
0.9	0.9016	1.109	0.45	0.9	0.5070	1.972	0.95	0.9	0.2851	3.508	0.45	0.9	0.1603	6.237	0.95
1.0	**0.8913**	**1.122**	**0.50**	**6.0**	**0.5012**	**1.995**	**3.00**	**11.0**	**0.2818**	**3.548**	**0.50**	**16.0**	**0.1585**	**6.310**	**8.00**
0.1	0.8810	1.135	0.55	0.1	0.4955	2.018	0.05	01	0.2786	3.589	0.55	01	0.1567	6.383	0.05
0.2	0.8710	1.148	0.60	0.2	0.4898	2.042	0.10	02	0.2754	3.631	0.60	02	0.1549	6.457	0.10
0.3	0.8610	1.161	0.65	0.3	0.4842	2.065	0.15	03	0.2723	3.673	0.65	03	0.1531	6.531	0.15
0.4	0.8511	1.175	0.70	0.4	0.4786	2.089	0.20	04	0.2692	3.715	0.70	04	0.1514	6.607	0.20
0.5	0.8414	1.189	0.75	0.5	0.4732	2.113	0.25	05	0.2661	3.758	0.75	05	0.1496	6.683	0.25
0.6	0.8318	1.202	0.80	0.6	0.4677	2.138	0.30	06	0.2630	3.802	0.80	06	0.1479	6.761	0.30
0.7	0.8222	1.216	0.85	0.7	0.4624	2.163	0.35	07	0.2600	3.846	0.85	07	0.1462	6.839	0.35
0.8	0.8128	1.230	0.90	0.8	0.4571	2.188	0.40	08	0.2570	3.890	0.90	08	0.1445	6.918	0.40
0.9	0.8035	1.245	0.95	0.9	0.4519	2.213	0.45	09	0.2541	3.936	0.95	09	0.1429	6.998	0.45
2.0	**0.7943**	**1.259**	**1.00**	**7.0**	**0.4467**	**2.239**	**0.50**	**12.0**	**0.2512**	**3.981**	**6.00**	**17.0**	**0.1413**	**7.079**	**0.50**
0.1	0.7852	1.274	0.05	0.1	0.4416	2.265	0.55	0.1	0.2483	4.027	0.05	0.1	0.1396	7.161	0.55
0.2	0.7762	1.288	0.10	0.2	0.4365	2.291	0.60	0.2	0.2455	4.074	0.10	0.2	0.1380	7.244	0.60
0.3	0.7674	1.303	0.15	0.3	0.4315	2.317	0.65	0.3	0.2427	4.121	0.15	0.3	0.1365	7.328	0.65
0.4	0.7586	1.318	0.20	0.4	0.4266	2.344	0.70	0.4	0.2399	4.169	0.20	0.4	0.1349	7.413	0.70
0.5	0.7499	1.334	0.25	0.5	0.4217	2.371	0.75	0.5	0.2371	4.217	0.25	0.5	0.1334	7.499	0.75
0.6	0.7413	1.349	0.30	0.6	0.4169	2.399	0.80	0.6	0.2344	4.266	0.30	0.6	0.1318	7.586	0.80
0.7	0.7328	1.365	0.35	0.7	0.4121	2.427	0.85	0.7	0.2317	4.315	0.35	0.7	0.1303	7.674	0.85
0.8	0.7244	1.380	0.40	0.8.	0.4074	2.455	0.90	0.8	0.2291	4.365	0.40	0.8	0.1288	7.762	0.90
0.9	0.7161	1.396	0.45	0.9	0.4027	2.483	0.95	0.9	0.2265	4.416	0.45	0.9	0.1274	7.852	0.95
3.0	**0.7079**	**1.413**	**0.50**	**8.0**	**0.3981**	**2.512**	**4.00**	**13.0**	**0.2239**	**4.467**	**0.50**	**18.0**	**0.1259**	**7.943**	**9.00**
0.1	0.6998	1.429	0.55	0.1	0.3936	2.541	0.05	0.1	0.2213	4.519	0.55	0.1	0.1245	8.035	0.05
0.2	0.6918	1.445	0.60	0.2.	0.3890	2.570	0.10	0.2	0.2188	4.571	0.60	0.2	0.1230	8.128	0.10
0.3	0.6839	1.462	0.65	0.3	0.3846	2.600	0.15	0.3	0.2163	4.624	0.65	0.3	0.1216	8.222	0.15
0.4	0.6761	1.479	0.70	0.4	0.3802	2.630	0.20	0.4	0.2138	4.677	0.70	0.4	0.1202	8.318	0.20
0.5	0.6683	1.496	0.75	0.5	0.3758	2.661	0.25	0.5	0.2113	4.732	0.75	0.5	0.1189	8.414	0.25
0.6	0.6607	1.514	0.80	0.6	0.3715	2.692	0.30	0.6	0.2089	4.786	0.80	0.6	0.1175	8.511	0.30
0.7	0.6531	1.531	0.85	0.7	0.3673	2.723	0.35	0.7	0.2065	4.842	0.85	0.7	0.1161	8.610	0.35
0.8	0.6457	1.549	0.90	0.8	0.3631	2.754	0.40	0.8	0.2042	4.898	0.90	0.8	0.1148	8.710	0.40
0.9	0.6383	1.567	0.95	0.9	0.3589	2.786	0.45	0.9	0.2018	4.955	0.95	0.9	0.1135	8.810	0.45
4.0	**0.6310**	**1.585**	**2.00**	**9.0**	**0.3548**	**2.818**	**0.50**	**1.4.0**	**0.1995**	**5.012**	**7.00**	**19.0**	**0.1122**	**8.913**	**0.50**
0.1	0.6237	1.603	0.05	0.1	0.3508	2.851	0.55	0.1	0.1972	5.070	0.05	0.1	0.1109	9.016	0.55
0.2	0.6166	1.622	0.10	0.2	0.3467	2.884	0.60	0.2	0.1950	5.129	0.10	0.2	0.1096	9.120	0.60
0.3	0.6095	1.641	0.15	0.3	0.3428	2.917	0.65	0.3	0.1928	5.188	0.15	0.3	0.1084	9.226	0.65

Table 48-3. Relationships between Decibel, Current, Voltage, and Power Ratios (Continued)

dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power
0.4	0.6026	1.660	0.20	0.4	0.3388	2.951	0.70	0.4	0.1905	5.248	0.20	0.4	0.1072	9.333	0.70
0.5	0. 5957	1. 679	0.25	0.5	0.3350	2.985	0.75	0.5	0.1884	5.309	0.25	0.5	0.1059	9.441	0.75
0.6	0.5888	1.698	0.30	0.6	0.3311	3.020	0.80	0.6	0.1862	5.370	0.30	0.6	0.1047	9.550	0.80
0.7	0. 5821	1.718	0.35	0.7	0.3273	3.055	0.85	0.7	0.1841	5.433	0.35	0.7	0.1035	9.661	0.85
0.8	0. 5754	1.738	0.40	0.8	0.3236	3.090	0.90	0.8	0.1820	5.495	0.40	0.8	0.1023	9.772	0.90
0.9	0.5689	1.758	0.45	0.9	0.3199	3.126	0.95	0.9	0.1799	5.559	0.45	0.9	0.1012	9.886	0.95

dB Voltage	Loss	Gain	dB Power	dB Voltage	Loss	Gain	dB Power
20.0	**0.1000**	**10.00**	**10.00**	**60.0**	**0.001**	**1,000**	**30.00**
	Use the same number as 0–20 dB but shift decimal point one step to the left. Thus since 10 dB = 0.3162 30 dB = 0.03162	Use the same number as 0–20 dB but shift decimal point one step to the right. Thus since 10 dB = 3.162 30 dB = 31.62	This column repeats every 10 dB instead of 20 dB		Use the same numbers as 0–20 dB but shift point three steps to the left. Thus since 10 dB = 0.3162 70 dB = 0.0003162	Use the same number as 0–20 dB column but shift point three steps to the right. Thus since 10 dB = 3.162 70 dB = 3162	This column repeats every 10 dB instead of 20 dB
40.0	**0.01**	**100**	**20**	**80**	**0.0001**	**10,000**	**40.00**
	Use the same number as 0–20 dB but shift point two steps to the left. Thus since 10 dB = 0 3162 50 dB = 0.003162	Use the same number as 0–20 dB but shift point two steps to the right. Thus since 10 dB = 3162 50 dB = 316.2	This column repeats every 10 dB instead of 20 dB		Use the same numbers as 0–20 dB but shift point four steps to the left. Thus since 10 dB = 0.3162 90 dB = 0.00003162	Use the same number as 0–20 dB but shift point four steps to the right. Thus since 10 dB = 3.162 90 dB = 31620	This column repeats every 10 dB instead of 20 dB
				100	**0.00001**	**100,000**	**50.00**

$$\text{dBm at new } Z = \text{dBm}_{600\Omega} + 10\log\frac{600\ \Omega}{Z_{new}} \tag{48-41}$$

Example*:* The dBm for a 32 Ω load is

$$\begin{aligned}\text{dBm}_{32} &= 4\ \text{dBm} + 10\log\frac{600\ \Omega}{32\ \Omega}\\ &= 16.75\ \text{dBm}\end{aligned}$$

This can also be determined by using the graph in Fig. 48-5.

To find the logarithm of a number to some other base than the base 10 and 2.718, use

$$n = b^L \tag{48-42}$$

A number is equal to a base raised to its logarithm,

$$\ln(n) = \ln(bL) \tag{48-43}$$

therefore,

$$\frac{\ln(n)}{\ln(b)} = L \tag{48-44}$$

The natural log is a number divided by the natural log of the base equals the logarithm.

Example: Find the logarithm of the number 2 to the base 10:

$$\begin{aligned}\frac{\ln 2}{\ln 10} &= \frac{0.693147}{2.302585}\\ &= 0.301030\end{aligned}$$

In information theory work, logarithms to the base 2 are quite commonly employed. To find the $\log_2$ of 26

$$\frac{\ln 26}{\ln 2} = 4.70$$

To prove this, raise 2 to the 4.70 power

$$2^{4.70} = 26$$

48.4 Sound Pressure Level

The *sound pressure level* (SPL) is related to acoustic pressure as seen in Fig. 48-6.

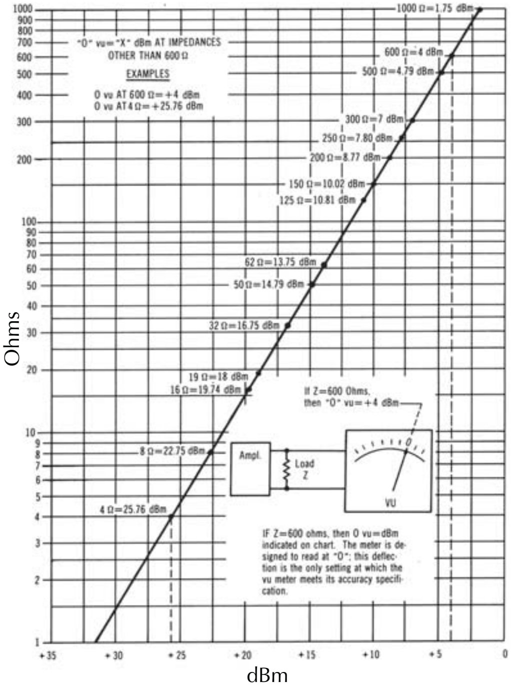


Figure 48-5. Relationship between VU and dBm at various impedances.

48.5 Sound System Quantities and Design Formulas

Various quantities used for sound system design are defined as follows:

***D_1*.** D_1 is the distance between the microphone and the loudspeaker, Fig. 48-7.

***D_2*.** D_2 is the distance between the loudspeaker and the farthest listener, Fig. 48-7.

***D_o*.** D_o is the distance between the talker (sound source) and the farthest listener, Fig. 48-7.

***D_s*.** D_s is the distance between the talker (sound source) and the microphone, Fig. 48-7.

***D_L*.** D_L is the *limiting distance* and is equal to 3.16 D_c for 15%*Alcons* in a room with a reverberation time of 1.6 s. This means that D_2 cannot be any longer than D_L if *Alcons* is to be kept at 15% or less. As the RT_{60} increases or the required $\%Alcon_s$ decreases D_2 becomes less than D_L.

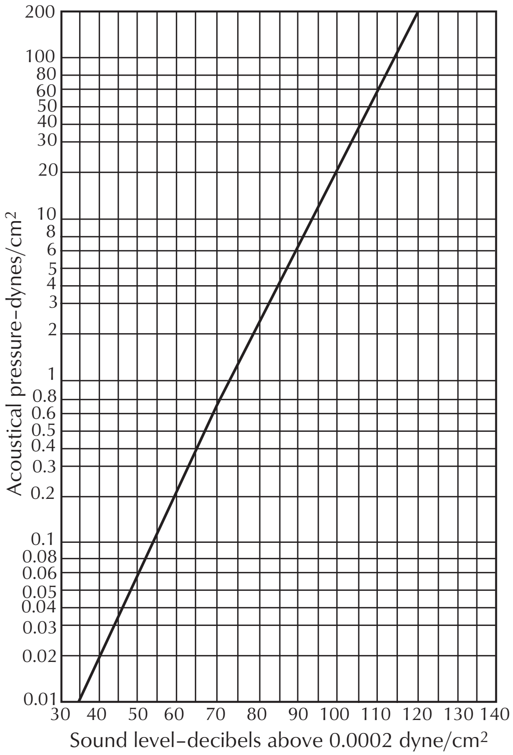


Figure 48-6. Sound pressure level versus acoustic pressure.

***EAD*.** The *equivalent acoustic distance* (*EAD*) is the maximum distance from the talker that produces adequate loudness of the unamplified voice. Often an *EAD* of 8 ft is used in quiet surroundings as it is the distance at which communications can be understood comfortably. Once the *EAD* has been determined, the sound system is designed to produce that level at every seat in the audience.

***D_c*.** *Critical distance* (D_c) is the point in a room where the direct sound and reverberant sound are equal. D_c is found by the equation

$$D_c = 0.141\frac{QRM}{N} \tag{48-45}$$

where,

Q is the directivity of the sound source,

R is the room constant,

M is the critical distance modifier for absorption coefficient,

N is the modifier for direct-to-reverberant speaker coverage.

It can also be found with the equation

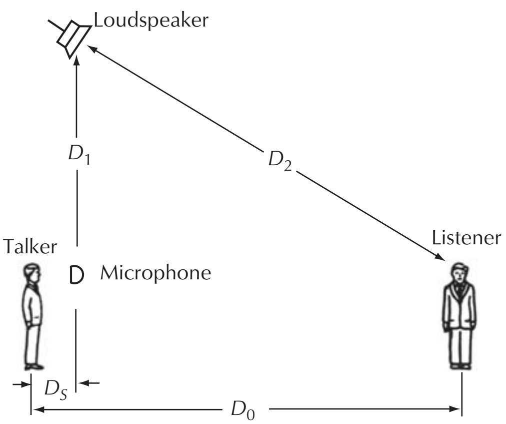


Figure 48-7. Definitions of sound system dimensions.

$$D_c = 0.3121^{**}\sqrt{\frac{Qv}{RT_{60}}} \tag{48-46}$$

** 0.057 for SI units

M. The critical distance modifier (*M*) corrects for the effect of a different absorption coefficient within the path of the loudspeaker's coverage pattern.

$$M = \frac{1-\bar{a}_{total\ room}}{1-\bar{a}_{loudspeaker\ coverage\ area}} \tag{48-47}$$

N. The critical distance modifier (*N*) corrects for multiple sound sources. *N* is the number describing the ratio of acoustic power going to the reverberant sound field without supplying direct sound versus the acoustic power going from the loudspeakers providing direct sound to a given listener position.

$$N = \frac{\text{Total number of loudspeakers}}{\text{Number providing direct sound}} \tag{48-48}$$

%Alcons. The English language is made up of consonants and vowels. The consonants are the harsh letters that determine words. If the consonants of words are understood, the sentences or phrases will be understood. V. M. A. Peutz and W. Klein of Holland developed and published equations for the % articulation loss of consonants (*%Alcons*). The equation is

$$\%Alcons = 656^{**}\frac{RT_{60}{}^2 D_2{}^2 N}{VQM} \tag{48-49}$$

** 200 for SI units

where,

Q is the directivity of the sound source,

V is the volume of the enclosure,

M is the critical distance modifier for absorption,

N is the critical distance modifier for multiple sources,

D_2 is the distance between the loudspeaker and the farthest listener.

When $D_c \geq D_L$, then $\%ALcons = 9RT_{60}$.

FSM. The *feedback stability margin* (*FSM*) is required to insure that a sound reinforcement system will not ring. A room and sound system, when approaching feedback, gives the effect of an long reverberation time. A room, for instance, with an RT_{60} of 3 s could easily have an apparent RT_{60} of 6–12 s when the sound system approaches feedback. To insure that this long reverberation time does not happen, a feedback stability margin of 6 dB is added into the needed acoustic gain equation.

NOM. The *number of open microphones* (*NOM*) affects the gain of a sound reinforcement system. The system gain will be reduced by the following equation:

$$Gain\ reduction_{dB} = 10\log NOM \tag{48-50}$$

Every time the number of microphones doubles, the gain from the previous microphones is halved as the total gain is the gain of all the microphones added together.

NAG. The *needed acoustic gain* (*NAG*) is required to produce the same level at the farthest listener as at the *EAD*. *NAG* in its simplest form is

$$NAG = 20\log D_o - 20\log EAD \tag{48-51}$$

NAG, however, is also affected by the number of open microphones (*NOM*) in the system. Each time the *NOM* doubles, the *NAG* increases 3 dB. Finally, a 6 dB feedback stability margin (*FSM*) is added into the *NAG* formula to ensure that the system never approaches feedback. The final equation for *NAG* is

$$NAG = \Delta D_o - \Delta EAD + 10\log NOM + 6\text{ dB FSM} \tag{48-52}$$

where,

ΔD_o and ΔEAD are the level change per the Hopkins-Stryker equation.

PAG. The *potential acoustic gain* (*PAG*) of a sound system is

$$PAG = \Delta D_o + \Delta D_1 - \Delta D_s - \Delta D_2 \tag{48-53}$$

where,

ΔD_o, ΔD_1, ΔD_s, and ΔD_2 are found as in *NAG.*

Q. The *directivity factor* (Q) of a transducer used for sound emission is the ratio of sound pressure squared, at some fixed distance and specified direction, to the mean sound pressure squared at the same distance averaged over all directions from the transducer. The distance must be great enough so that the sound appears to diverge spherically from the effective acoustic center of the source. Unless otherwise specified, the reference direction is understood to be that of maximum response.

Geometric Q can be found by using the following equations:

1. For rectangular coverage between 0 degrees and 180 degrees,

$$Q_{Geom} = \frac{180}{\arcsin\left(\frac{\sin\theta}{2}\right)\left(\frac{\sin\phi}{2}\right)} \tag{48-54}$$

2. For angles between 180 degrees and 360 degrees when one angle is 180 degrees, and the other angle is some value between 180 degrees and 360 degrees

$$Q_{Geom} = \frac{360}{\text{angle}} \tag{48-55}$$

3. For conical coverage,

$$Q_{Geom} = \frac{2}{1 - \cos\frac{\theta}{2}} \tag{48-56}$$

$C_\angle$. $C_\angle$ is the *included angle* of the coverage pattern. Normally $C_\angle$ is expressed as an angle between the –6 dB points in the coverage pattern.

EPR. *EPR* is the *electrical power required* to produce the desired SPL at a specific point in the coverage area. It is found by the equation

$$EPR_{watts} = 10^{\frac{\text{SPL}_{des} + 10\text{dB}_{crest} + \Delta D_2 - \Delta D_{ref} - L_{sens}}{10}} \tag{48-57}$$

a. The *absorption coefficient* (a) of a material or surface is the ratio of absorbed sound to reflected sound or incident sound

$$a = \frac{I_A}{I_R} \tag{48-58}$$

If all sound was reflected, a would be 0. If all sound were absorbed, a would be 1.

***$\bar{a}$*.** The *average absorption coefficient* ($\bar{a}$) for all the surfaces together and is found by

$$\bar{a} = \frac{S_1 a_1 + S_2 a_2 + \ldots\ S_n a_n}{S} \tag{48-59}$$

where,
$S_{1,2\ \ldots\ n}$ are individual surface areas,
$a_{1,2\ \ldots\ n}$ are the individual absorption coefficients of the areas,
S is the total surface area.

MFP. The *mean-free path* (*MFP*) is the average distance between reflections in a space. *MFP* is found by

$$MFP = \frac{4V}{S} \tag{48-60}$$

where,
V is the space volume,
S is the space surface area.

ΔD_x. ΔD_x is an arbitrary level change associated with the specific distance from the Hopkins-Stryker equation so that

$$\Delta D_x = -10\log\left[\frac{Q}{4\pi {D_x}^2} + \frac{4N}{Sa}\right] \tag{48-61}$$

In semireverberant rooms, Peutz describes ΔD_x as

$$\Delta D_x = -10\log\left[\frac{Q}{4\pi {D_x}^2} + \frac{4N}{Sa}\right] + \frac{0.734^{**}\sqrt{V}}{h\text{RT}_{60}}\log\frac{D_x > D_\text{c}}{D_c} \tag{48-62}$$

** 200 for SI units

where,
h is the ceiling height.

SNR. *SNR* is the acoustical *signal-to-noise ratio*. The signal-to-noise ratio required for intelligibility is

$$SNR = 35\left(\frac{2 - \log\%Alcons}{2 - \log 9RT_{60}}\right) \tag{48-63}$$

SPL. *SPL* is the *sound pressure level* in dB-SPL re 0.00002 N/m². SPL is also called L_p.

Max Program Level. *Max program level* is the maximum program level attainable at a specific point from the available input power. Max program level is

$$program\ level_{max} = 10\log\frac{watts_{avail}}{10} - (\Delta D_2 - \Delta D_{ref}) + L_{sens} \tag{48-64}$$

***L*$_{sens}$.** *Loudspeaker sensitivity* (L_{sens}) is the on-axis SPL output of the loudspeaker with a specified power input and at a specified distance. The most common L_{sens} are at 4 ft, 1 W and 1 m, and 1 W.

Sa. *Sa* is the *total absorption* in sabines of all the surface areas times their absorption.

***dB-SPL*$_T$.** The $dB\text{-}SPL_T$ is the talker's or sound source's *sound pressure level*.

***dB-SPL*$_D$.** The $dB\text{-}SPL_D$ is the *desired sound pressure level*.

dB-SPL. The dB-SPL is the *sound pressure level* in decibels.

EIN. *EIN* is the *equivalent input noise*.

$$EIN = -198\text{ dB} + 10\log BW + 10\log Z - 6\text{ dB} - 20\log 0.775 \tag{48-65}$$

where,

BW is the bandwidth,

Z is the impedance.

Thermal Noise. *Thermal noise* is the noise produced in any resistance, including standard resistors. Any resistance that is at a temperature above absolute zero generates noise due to the thermal agitation of free electrons in the material. The magnitude of the noise can be calculated from the resistance, absolute temperature, and equivalent noise bandwidth of the measuring system. A completely noise-free amplifier whose input is connected to its equivalent source resistance will have noise in its output equal to the product of amplification and source resistor noise. This noise is said to be the *theoretical minimum*.

Fig. 48-8 provides a quick means for determining the rms value of thermal noise voltage in terms of resistance and circuit bandwidth.

For practical calculations, especially those in which the resistive component is constant across the bandwidth of interest, use

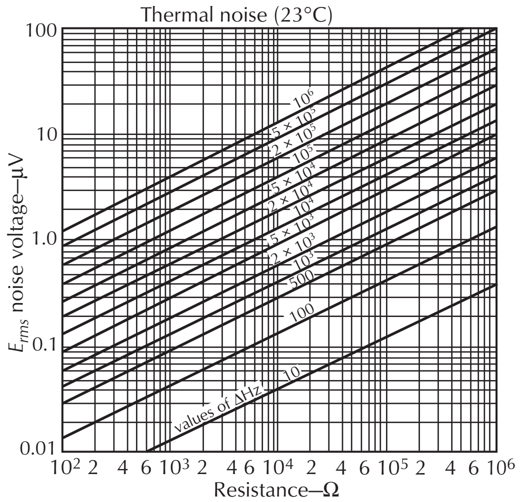


Figure 48-8. Thermal noise graph.

$$E_{rms} = \sqrt{4 \times 10^{-23}(T)(f_1 - f_2)R} \tag{48-66}$$

where,

$f_1 - f_2$ is the 3 dB bandwidth,

R is the resistive component of the impedance across which the noise is developed,

T is the absolute temperature in K.

***RT*$_{60}$.** RT_{60} is the time required for an interrupted steady-state signal in a space to decay 60 dB. RT_{60} is normally calculated using one of the following equations: the classic Sabine method, the Norris Eyring modification of the Sabine equation, and the Fitzroy equation. The Fitzroy equation is best used when the walls in the *X*, *Y*, and *Z* planes have very different absorption materials on them.

Sabine:

$$RT_{60} = 0.049^{**}\frac{V}{S\bar{a}} \tag{48-67}$$

** 0.161 for SI units

Norris Eyring:

$$RT_{60} = 0.049^{**}\frac{V}{-S\ln(1-\bar{a})} \tag{48-68}$$

** 0.161 for SI unit.

Fitzroy:

$$RT_{60} = \frac{0.049^{**}V}{S^2}\left[\frac{2XY}{-\ln(1-\bar{a}_{XY})} + \frac{2XZ}{-\ln(1-\bar{a}_{XZ})} + \frac{2YZ}{-\ln(1-\bar{a}_{YZ})}\right] \qquad (48\text{-}69)$$

** 0.161 for SI units

where,
V is the room volume,
S is the surface area,
a is the total absorption coefficient,
X is the space length,
Y is the space width,
Z is the space height.

Signal Delay. *Signal delay* is the time required for a signal, traveling at the speed of sound, to travel from the source to a specified point in space

$$SD = \frac{Distance}{c} \qquad (48\text{-}70)$$

where,
SD is the signal delay in milliseconds,
c is the speed of sound.

48.6 ISO Numbers

"Preferred Numbers were developed in France by Charles Renard in 1879 because of a need for a rational basis for grading cotton rope. The sizing system that resulted from his work was based upon a geometric series of mass per unit length such that every fifth step of the series increased the size of rope by a factor of ten." (From the American National Standards for Preferred Numbers). This same system of preferred numbers is used today in acoustics. The one-twelfth, one-sixth, one-third, one-half, two-thirds, and one octave preferred center frequency numbers are not the exact n series number. The exact n series number is found by the equation

$$n \text{ Series number} = 10^{\frac{1}{n}}\left(10^{\frac{1}{n}}\right)\left(10^{\frac{1}{n}}\right)\ldots \qquad (48\text{-}71)$$

where,
n is the ordinal numbers in the series.

For instance, the third n number for a 40 series would be

$$10^{\frac{1}{40}}\left(10^{\frac{1}{40}}\right)\left(10^{\frac{1}{40}}\right) = 1.1885022$$

The preferred ISO number is 1.18. Table 48-4 is a table of preferred International Standards Organization (ISO) numbers.

Table 48-4. Internationally Preferred ISO Numbers

⅟12 oct.	⅙ oct.	⅓ oct.	½ oct.	⅔ oct.	⅟1 oct.	
40 ser.	**20 ser.**	**10 ser.**	**6⅓ ser.**	**5 ser.**	**3⅓ ser.**	**Exact value**
1.00	1.00	1.00	1.00	1.00	1.00	1.000000000
1.06						1.059253725
1.12	1.12					1.122018454
1.18						1.188502227
1.25	1.25	1.25				1.258925411
1.32						1.333521431
1.40	1.40					1.412537543
1.50						1.496235654
1.60	1.60	1.60		1.60		1.584893190
1.70						1.678804015
1.80	1.80					1.778279406
1.90						1.883649085
2.00	2.00	2.00	2.00		2.00	1.995262310
2.12						2.113489034
2.24	2.24					2.238721132
2.36						2.371373698
2.50	2.50	2.50		2.50		2.511886423
2.65						2.660725050
2.80	2.80		2.80			2.818382920
3.00						2.985382606
3.15	3.15	3.15				3.162277646
3.35						3.349654376
3.55	3.55					3.548133875
3.75						3.758374024
4.00	4.00	4.00	4.00	4.00	4.00	3.981071685
4.25						4.216965012
4.50	4.50					4.466835897
4.75						4.731512563
5.00	5.00	5.00				5.011872307
5.30						5.308844410
5.60	5.60		5.60			5.623413217
6.00						5.956621397
6.30	6.30	6.30		6.30		6.309573403
6.70						6.683439130
7.10	7.10					7.079457794
7.50						7.498942039
8.00	8.00	8.00	8.00		8.00	7.943282288
8.50						8.413951352

Table 48-4. Internationally Preferred ISO Numbers (Continued)

1/12 oct.	1/6 oct.	1/3 oct.	1/2 oct.	2/3 oct.	1/1 oct.	
40 ser.	20 ser.	10 ser.	6⅓ ser.	5 ser.	3⅓ ser.	Exact value
9.00	9.00					8.912509312
9.50						9.440608688

48.7 Greek Alphabet

The *Greek alphabet* plays a major role in the language of engineering and sound. Table 48-5 shows the Greek alphabet and the terms that are commonly symbolized by it.

Table 48-5. Greek Alphabet

Name	Upper Case		Lower Case	
alpha	A		α	absorption factor, angles, angular acceleration, attenuation constant, common-base current amplification factor, deviation of state parameter, temperature coefficient of linear expansion, temperature coefficient of resistance, thermal expansion coefficient, thermal diffusivity
beta	B		ß	angles, common-emitter current amplification factor, flux density, phase constant, wavelength constant
gamma	Γ		γ	electrical conductivity, Grueneisen parameter
delta	Δ	decrement increment	δ	angles, damping coefficient (decay constant), decrement, increment, secondary-emission ratio
epsilon	E	electric field intensity	ε	capacitivity, dielectric coefficient, electron energy, emissivity, permittivity, base of natural logarithms (2.71828)
zeta	Z		ζ	chemical potential, dielectric susceptibility (intrinsic capacitance), efficiency, hysteresis, intrinsic impedance of a medium, intrinsic standoff ratio
eta	H		η	
theta	Θ	angles, thermal resistance	θ	angle of rotation, angles, angular phase displacement, reluctance, transit angle
iota	I		ι	
kappa	K	coupling coefficient	κ	susceptibility
lambda	Λ		λ	line density of charge, permeance, photosensitivity, wavelength

Table 48-5. Greek Alphabet (Continued)

Name	Upper Case		Lower Case	
mu	M		μ	amplification factor, magnetic permeability, micron, mobility, permeability, prefix micro
nu	N		ν	reluctivity
xi	Ξ		ξ	
omicron	O		o	
pi	Π		π	Peltier coefficient, ratio of circumference to diameter (3.1416)
rho	P		p	reflection coefficient, reflection factor, resistivity, volume density of electric charge
sigma	Σ	summation	σ	conductivity, Stefan-Boltzmann constant, surface density of charge
tau	T	period	τ	propagation constant, Thomson coefficient, time constant, time-phase displacement, transmission factor
upsilon	Y	admittance	υ	
phi	Φ	magnetic flux, radiant flux	ϕ	angles, coefficient of performance, contact potential, magnetic flux, phase angle, phase displacement
chi	X			angles
psi	Ψ	angles	ψ	dielectric flux, displacement flux, phase difference
omega	Ω	resistance	ω	angular frequency, angular velocity, solid angle

48.8 Audio Standards

Audio standards are defined by the and the Audio Engineering Society (AES), Table 48-6 and International Electrotechnical Commission (IEC) Table 48-7.

Table 48-6. AES Standards

Standards and Recommended Practices, Issued Jan 2007	
AES2-1984: (r2003)	AES recommended practice—Specification of loudspeaker components used in professional audio and sound reinforcement
AES3-2003:	AES recommended practice for digital audio engineering—Serial transmission format for two-channel linearly represented digital audio data (Revision of AES3-1992, including subsequent amendments)
AES5-2003:	AES recommended practice for professional digital audio—Preferred sampling frequencies for applications employing pulse-code modulation (revision of AES5-1997)
AES6-1982: (r2003)	Method for measurement of weighted peak flutter of sound recording and reproducing equipment

Table 48-6. AES Standards (Continued)

Standards and Recommended Practices, Issued Jan 2007	
AES7-2000: (r2005)	AES standard for the preservation and restoration of audio recording—Method of measuring recorded fluxivity of magnetic sound records at medium wavelengths (Revision of AES7-1982)
AES10-2003:	AES recommended practice for digital audio engineering—Serial Multichannel Audio Digital Interface (MADI) (Revision of AES10-1991)
AES11-2003:	AES recommended practice for digital audio engineering—Synchronization of digital audio equipment in studio operations. (Revision of AES11-1997)
AES14-1992: (r2004)	AES standard for professional audio equipment—Application of connectors, part 1, XLR-type polarity and gender
AES15-1991: (w2002)	AES recommended practice for sound-reinforcement systems—Communications interface (PA-422) (Withdrawn: 2002)
AES17-1998: (r2004)	AES standard method for digital audio engineering—Measurement of digital audio equipment (Revision of AES17-1991)
AES18-1996: (r2002)	AES recommended practice for digital audio engineering—Format for the user data channel of the AES digital audio interface. (Revision of AES18-1992)
AES19-1992: (w2003)	AES-ALMA standard test method for audio engineering—Measurement of the lowest resonance frequency of loudspeaker cones (Withdrawn: 2003)
AES20-1996: (r2002)	AES recommended practice for professional audio—Subjective evaluation of loudspeakers
AES22-1997: (r2003)	AES recommended practice for audio preservation and restoration—Storage and handling—Storage of polyester-base magnetic tape
AES24-1-1999: (w2004)	AES standard for sound system control—Application protocol for controlling and monitoring audio devices via digital data networks—Part 1: Principles, formats, and basic procedures (Revision of AES24-1-1995)
AES24-2-tu: (w2004)	PROPOSED DRAFT AES standard for sound system control—Application protocol for controlling and monitoring audio devices via digital data networks—Part 2, data types, constants, and class structure (for Trial Use)
AES26-2001:	AES recommended practice for professional audio—Conservation of the polarity of audio signals (Revision of AES26-1995)
AES27-1996: (r2002)	AES recommended practice for forensic purposes—Managing recorded audio materials intended for examination
AES28-1997: (r2003)	AES standard for audio preservation and restoration—Method for estimating life expectancy of compact discs (CD-ROM), based on effects of temperature and relative humidity (includes Amendment 1-2001)
AES31-1-2001: (r2006)	AES standard for network and file transfer of audio—Audio-file transfer and exchange Part 1: Disk format
AES31-2-2006:	AES standard on network and file transfer of audio—Audio-file transfer and exchange—File format for transferring digital audio data between systems of different type and manufacture
AES31-3-1999:	AES standard for network and file transfer of audio—Audio-file transfer and exchange—Part 3: Simple project interchange
AES32-tu:	PROPOSED DRAFT AES standard for professional audio interconnections—Fibre optic connectors, cables, and characteristics (for Trial Use)
AES33-1999: (w2004)	AES standard—For audio interconnections—Database of multiple—program connection configurations (Withdrawn: 2004)
AES35-2000: (r2005)	AES standard for audio preservation and restoration—Method for estimating life expectancy of magneto-optical (M-O) disks, based on effects of temperature and relative humidity
AES38-2000: (r2005)	AES standard for audio preservation and restoration—Life expectancy of information stored in recordable compact disc systems—Method for estimating, based on effects of temperature and relative humidity
AES41-2000: (r2005)	AES standard for digital audio—Recoding data set for audio bit-rate reduction
AES42-2006:	AES standard for acoustics—Digital interface for microphones
AES43-2000: (r2005)	AES standard for forensic purposes—Criteria for the authentication of analog audio tape recordings
AES45-2001:	AES standard for single program connectors—Connectors for loudspeaker-level patch panels
AES46-2002:	AES standard for network and file transfer of audio Audio-file transfer and exchange, Radio traffic audio delivery extension to the broadcast-WAVE-file format
AES47-2006:	AES standard for digital audio—Digital input-output interfacing—Transmission of digital audio over asynchronous transfer mode (ATM) networks
AES48-2005:	AES standard on interconnections—Grounding and EMC practices—Shields of connectors in audio equipment containing active circuitry
AES49-2005:	AES standard for audio preservation and restoration—Magnetic tape—Care and handling practices for extended usage
AES50-2005:	AES standard for digital audio engineering—High-resolution multichannel audio interconnection
AES51-2006:	AES standard for digital audio—Digital input-output interfacing—Transmission of ATM cells over Ethernet physical layer
AES52-2006:	AES standard for digital audio engineering—Insertion of unique identifiers into the AES3 transport stream
AES53-2006:	AES standard for digital audio—Digital input-output interfacing—Sample-accurate timing in AES47
Duplicate entry:	Duplicate entry
AES-1id-1991: (r2003)	AES information document—Plane wave tubes: design and practice
AES-2id-1996: (r2001)	AES information document for digital audio engineering—Guidelines for the use of the AES3 interface

Table 48-6. AES Standards (Continued)

Standards and Recommended Practices, Issued Jan 2007	
AES-3id-2001: (r2006)	AES information document for digital audio engineering—Transmission of AES3 formatted data by unbalanced coaxial cable (Revision of AES-3id-1995)
AES-4id-2001:	AES information document for room acoustics and sound reinforcement systems—Characterization and measurement of surface scattering uniformity
AES-5id-1997: (r2003)	AES information document for room acoustics and sound reinforcement systems—Loudspeaker modeling and measurement—Frequency and angular resolution for measuring, presenting, and predicting loudspeaker polar data
AES-6id-2000:	AES information document for digital audio—Personal computer audio quality measurements
AES-10id-2005:	AES information document for digital audio engineering—Engineering guidelines for the multichannel audio digital interface, AES10 (MADI)
AES-R1-1997:	AES project report for professional audio—Specifications for audio on high-capacity media
AES-R2-2004:	AES project report for articles on professional audio and for equipment specifications—Notations for expressing levels (Revision of AES-R2-1998)
AES-R3-2001:	AES standards project report on single program connector—Compatibility for patch panels of tip-ring-sleeve connectors
AES-R4-2002:	AES standards project report. Guidelines for AES Recommended practice for digital audio engineering—Transmission of digital audio over asynchronous transfer mode (ATM) networks
AES-R6-2005:	AES project report—Guidelines for AES standard for digital audio engineering—High-resolution multichannel audio interconnection (HRMAI)
AES-R7-2006:	AES standards project report—Considerations for accurate peak metering of digital audio signals

Table 48-7. IEC Standards

IEC Number	IEC Title
IEC 60038	IEC standard voltages
IEC 60063	Preferred number series for resistors and capacitors
IEC 60094	Magnetic tape sound recording and reproducing systems
IEC 60094-5	Electrical magnetic tape properties
IEC 60094-6	Reel-to-reel systems
IEC 60094-7	Cassette for commercial tape records and domestic use
IEC 60096	Radio-frequency cables
IEC 60098	Rumble measurement on vinyl disc turntables
IEC 60098	Rumble measurement on vinyl disc turntables
IEC 60134	Absolute maximum and design ratings of tube and semiconductor devices
IEC 60169	Radio-frequency connectors
IEC 60169-2	Unmatched coaxial connector (Belling-Lee TV Aerial Plug)
IEC 60169-8	BNC connector, 50 ohm
IEC 60169-9	SMC connector, 50 ohm
IEC 60169-10	SMB connector, 50 ohm
IEC 60169-15	N connector, 50 ohm or 75 ohm
IEC 60169-16	SMA connector, 50 ohm
IEC 60169-16	TNC connector, 50 ohm
IEC 60169-24	F connector, 75 ohm
IEC 60179	Sound level meters
IEC 60228	Conductors of insulated cables
IEC 60268	Sound system equipment
IEC 60268-1	General
IEC 60268-2	Explanation of general terms and calculation methods
IEC 60268-3	Amplifiers
IEC 60268-4	Microphones
IEC 60268-5	Loudspeakers
IEC 60268-6	Auxiliary passive elements
IEC 60268-7	Headphones and earphones
IEC 60268-8	Automatic gain control devices
IEC 60268-9	Artificial reverberation, time delay, and frequency shift equipment
IEC 60268-10	Peak program level meters
IEC 60268-11	Application of connectors for the interconnection of sound system components
IEC 60268-12	Application of connectors for broadcast and similar use
IEC 60268-13	Listening tests on loudspeakers
IEC 60268-14	Circular and elliptical loudspeakers; outer frame diameters and mounting dimensions
IEC 60268-16	Objective rating of speech intelligibility by speech transmission index
IEC 60268-17	Standard volume indicators
IEC 60268-18	Peak program level meters—Digital audio peak level meter
IEC 60297	19-inch rack
IEC 60386	Wow and flutter measurement (audio)
IEC 60417	Graphical symbols for use on equipment
IEC 60446	Wiring colors
IEC 60574	Audio-visual, video, and television equipment and systems
IEC 60651	Sound level meters
IEC 60908	Compact disk digital audio system
IEC 61043	Sound intensity meters with pairs of microphones
IEC 61603	Infrared transmission of audio or video signals
IEC 61966	Multimedia systems—Color measurement
IEC 61966-2-1	sRGB default RGB color space

48.9 Audio Frequency Range

The audio spectrum is usually considered the frequency range between 20 Hz and 20 kHz, Fig. 48-10. In reality, the upper limit of hearing pure tones is between 12 kHz and 18 kHz, depending on the person's age and sex and how well the ears have been trained and protected against loud sounds. Frequencies above 20 kHz cannot be heard as a sound, but the effect created by such frequencies (i.e., rapid rise time) can be heard.

The lower end of the spectrum is more often felt than heard as a pure tone. Frequencies below 20 Hz are difficult to reproduce. Often the reproducer actually reproduces the second harmonic of the frequency, and the brain translates it back to the fundamental.

48.10 Common Conversion Factors

Conversion from U.S. to SI units can be made by multiplying the U.S. unit by the conversion factors in Table 48-8. To convert from SI units to U.S. units, divide by the conversion factor.

Table 48-8. U.S. to SI Units Conversion Factors

U.S. Unit	Multiplier	SI Unit
Length		
ft	3.048 000 × 10^{-1}	m
mi	1.609 344 × 10^{3}	m
in	2.540 000 × 10^{-2}	m
Area		
ft^2	9.290 304 × 10^{-2}	m^2
in^2	6.451 600 × 10^{-4}	m^2
yd^2	8.361 274 × 10^{-1}	m^2
Capacity/volume		
in^3	1.638 706 × 10^{-5}	m^3
ft^3	2.831 685 × 10^{-2}	m^3
liquid gal	3.785 412 × 10^{-3}	m^3
Volume/mass		
ft^3/lb	6.242 796 × 10^{-2}	m^3/kg
in^3/lb	3.612 728 × 10^{-5}	m^3/kg
Velocity		
ft/h	4.466 667 × 10^{-5}	m/s
in/s	2.540 000 × 10^{-2}	m/s
mi/h	4.470 400 × 10^{-1}	m/s
Mass		
oz	2.834 952 × 10^{-2}	kg
lb	4.535 924 × 10^{-1}	kg
Short Ton (2000 lb)	9.071 847 × 10^{2}	kg
Long Ton (2240 lb)	1.016 047 × 10^{3}	kg
Mass/volume		
oz/in^3	1.729 994 × 10^{3}	kg/m^3
lb/ft^3	1.601 846 × 10^{1}	kg/m^3
lb/in^3	2.767 990 × 10^{4}	kg/m^3
lb/U.S. Gal	1.198 264 × 10^{2}	kg/m^3
Acceleration		
ft/s^2	3.048 000 × 10^{-1}	m/s^2
Angular Momentum		
lb f^2/s	4.214 011 × 10^{-2}	kg·m^2/s
Electricity		
A•h	3.600 000 × 10^{3}	C
Gs	1.000 000 × 10^{-4}	T
Mx	1.000 000 × 10^{-8}	Wb
Mho	1.000 000 × 10^{0}	S
Oe	7.957 747 × 10^{1}	A/m
Energy (Work)		
Btu	1.055 056 × 10^{3}	J
eV	1.602 190 × 10^{-19}	J
W•h	3.600 000 × 10^{3}	J
erg	1.000 000 × 10^{-7}	J
Cal	4.186 800 × 10^{0}	J
Force		
dyn	1.000 000 × 10^{-5}	N
lbf	4.448 222 × 10^{0}	N
pdl	1.382 550 × 10^{-1}	N
Heat		
Btu/ft^2	1.135 653 × 10^{4}	J/m^2
Btu/lb	2.326 000 × 10^{3}	J/hg
Btu/(h•ft^2•°F) or k (thermal conductivity)	1.730 735 × 10^{0}	W/m•K
Btu/(h•f^2•°F) or C (thermal conductance)	5.678 263 × 10^{0}	W/m^2•K
Btu/(lb °F) or c (heat capacity)	4.186 800 × 10^{3}	J/kg•K
°F•h•ft^2/Btu or R (thermal resistance)	1.761 102 × 10^{-1}	K•m^2/W
cal	4.186 000 × 10^{0}	J
cal/g	4.186 000 × 10^{3}	J/kg
Light		
Cd (candle power)	1.000 000 × 10^{0}	cd (candela)
fc	1.076 391 × 10^{1}	lx
fL	3.426 259 × 10^{0}	cd/m^2
Moment of Inertia		
lb•ft^2	4.214 011 × 10^{-2}	kg•m^2
Momentum		
lb•ft/s	1.382 550 × 10^{-1}	kg•m/s
Power		
Btu/h	2.930 711 × 10^{-1}	W
erg/s	1.000 000 × 10^{-7}	W

Table 48-8. U.S. to SI Units Conversion Factors

U.S. Unit	Multiplier	SI Unit
hp (550 ft/lb/s)	$7.456\,999 \times 10^{2}$	W
hp (electric)	$7.460\,000 \times 10^{2}$	W
Pressure		
atm (normal atmosphere)	$1.031\,250 \times 10^{5}$	Pa
bar	$1.000\,000 \times 10^{5}$	Pa
in Hg@ 60°F	$3.376\,850 \times 10^{3}$	Pa
dyn/cm^2	$1.000\,000 \times 10^{-1}$	Pa
cm Hg@ 0°C	$1.333\,220 \times 10^{3}$	Pa
lbf/f^2	$4.788\,026 \times 10^{1}$	Pa
pdl/ft^2	$1.488\,164 \times 10^{0}$	Pa
Viscosity		
cP	$1.000\,000 \times 10^{-3}$	Pa•s
lb/ft•s	$1.488\,164 \times 10^{0}$	Pa•s
ft^2/s	$9.290\,304 \times 10^{-2}$	m^2/s
Temperature		
°C	$t_C + 273.15$	K
°F	$(t_F + 459.67)/1.8$	K
°R	$t_R/1.8$	K
°F	$(t_F - 32)/1.8$ or $(t_F - 32) \times (5/9)$	°C
°C	$1.8\,(t_C) + 32$ or $(t_C \times (9/5)) + 32$	°F

48.11 Technical Abbreviations

Many units or terms in engineering have abbreviations accepted either by the U.S. government or by the acousticians and audio consultants and engineers. Table 48-9 is a list of many of these abbreviations. Symbols for multiple and submultiple prefixes are shown in Table 48-1.

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
1000 electron volts	keV
A-weighted sound-pressure level in decibels	dBA
absorption coefficient	a
ac current	Iac
ac volt	Vac
acoustic intensity	I_a
Acoustical Society of America	ASA
adaptive delta pulse code modulation	ADPCM
admittance	Y
advanced access control system	AACS
advanced audio coding	AAC
advanced encryption standard	AES
advanced technology attachment\	ATA
Advanced Television Systems Committee	ATSC
alien crosstalk margin computation	ACMC
alien far-end crosstalk	AFEXT
alien near-end crosstalk	ANEXT
all-pass filter	APF
alternating current	ac
aluminum steel polyethylene	ASP
ambient noise level	ANL
American Broadcasting Company	ABC
American Federation of Television and Radio Artists	AFTRA
American National Standards Institute	ANSI
American Society for Testing and Materials	ASTM
American Society of Heating, Refrigeration and Air Conditioning Engineers	ASHRAE
American Standard Code for Information Interchange	ASCII
American Standards Association	ASA
American wire gauge	AWG
Americans with Disabilities Act	ADA
Americans with Disabilities Act	ADA
ampere	A
ampere-hour	Ah
ampere-turn	At
amplification factor	μ
amplitude modulation	AM
analog to digital	A/D
analog-to-digital converter	ADC
angstrom	Å
antilogarithm	antilog
apple file protocol	AFP
appliance wiring material	AWM
articulation index	AI
assisted resonance	AR
assistive listening devices	ALD
assistive listening systems	ALS
asymmetric digital subscriber line	ADSL
asynchronous transfer mode	ATM
atmosphere normal atmosphere technical atmosphere	atm at
atomic mass unit (unified)	u
attenuation crosstalk ratio	ACR
attenuation-to-cross talk ratio	ACR
Audio Engineering Society	AES
audio erase	AE
audio frequency	AF

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
audio high density	AHD
audio over IP	AoIP
audio/video receivers	AVR
automated test equipment	ATE
automatic frequency control	AFC
automatic gain control	AGC
automatic level control	ALC
automatic volume control	AVC
auxiliary	aux
available bit rate	ABR
available input power	AIP
avalanche photodiodes	APD
average	avg
average absorption coefficient	a
average amplitude	A_{avg}
average power	P_{avg}
backlight compensation	BLC
backward-wave oscillator	BWO
balanced current amplifier	BCA
balanced to unbalanced (Bal-Un)	Balun
bandpass filter	BPF
bandpass in hertz	f_{BP}
bandwidth	BW
bar	bar
barn	b
basic rate interface ISDN	BRI
baud	Bd
Bayonet Neil-Concelman	BNC
beat-frequency oscillator	BFO
bel	B
binary coded decimal	BCD
binary phase shift keying	BPSK
binaural impulse response	BIR
bipolar junction transistor	BJT
bit	b
bit error rate	BER
bits per second	bps
blue minus luminance	B-Y
breakdown voltage	BV
British Standards Institution	BSI
British thermal unit	Btu
building automation systems	BAS
bulletin board service	BBS
Butadiene-acrylonitrile copolymer rubber	NBR
calorie (International Table calorie)	cal_{IT}
calorie (thermochemical calorie)	cal_{th}
Canadian Electrical Code	CEC
Canadian Standards Association	CSA
candela	cd
candela per square foot	cd/ft^2
candela per square meter	cd/m^2
candle	cd
capacitance; capacitor	C
capacitive reactance	X_C
carrier sense multiple access/collision detection	CSMA/CD
carrier-sense multiple access with collision detection	CSMA/CD
carrierless amplitude phase modulation	CAP
cathode-ray oscilloscope	CRO
cathode-ray tube	CRT
cd universal device format	CD-UDF
centimeter	cm
centimeter-gram-second	CGS
central office	CO
central processing unit	CPU
certified technology specialist	CTS
charge coupled device	CCD
charge transfer device	CTD
chlorinated polyethylene	CPE
circular mil	cmil
citizens band	CB
closed circuit television	CCTV
coated aluminum polyethylene basic sheath	Alpeth
coated aluminum, coated steel	CASPIC
coated aluminum, coated steel, polyethylene	CACSP
coercive force	H_c
Columbia Broadcasting Company	CBS
Comité Consultatif International des Radiocommunications	CCIR
commercial online service	COLS
Commission Internationale de l'Eclairage	CIE
common mode rejection or common mode rejection ratio	CMR, CMRR
Communications Cable and Connectivity Cable Association	CCCA
compact disc	CD
compact disc digital audio	CD-DA
compact disc interactive	CD-I
compression/decompression algorithm	CODEC
computer aided design	CAD
conductor flat cable	FCFC
consolidation point	CP
constant angular velocity	CAV

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
constant bandwidth	CB
constant bandwidth filter	CBF
constant bit rate	CBR
constant linear velocity	CLV
constant percentage bandwidth	CPB
constant-amplitude phase-shift	CAPS
Consumer Electronics Association	CEA
contact resistance stability	CRS
content scrambling system	CSS
Continental Automated Building Association	CABA
continuous wave	CW
coulomb	C
coverage angle	C
critical bands	CB
critical distance	D_c
Cross Interleave Reed Solomon Code	CIRC
crosslinked polyethylene	XLPE
cubic centimeter	cm^3
cubic foot	ft^3
cubic foot per minute	ft^3/min
cubic foot per second	ft^3/s
cubic inch	in^3
cubic meter	m^3
cubic meter per second	m^3/s
cubic yard	yd^3
curie	Ci
Custom Electronics Design and Installation Association	CEDIA
customer service representative	CSR
cycle per second	Hz
cyclic redundancy check	CRC
data encryption standard	DES
data over cable service interface	DOCSIS
dc current	I_{dc}
dc voltage	V_{dc}
decibel	dB
decibel ref to one milliwatt	dBm
decibels with a reference of 1 V	dBV
deferred procedure calls	DPC
degree (plane angle)	...°
degree (temperature)	
degree Celsius	°C
degree Fahrenheit	°F
denial of service	DoS
dense wave division multiplexing	DWDM
depth of discharge	DOD
descriptive video service	DVS

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
Deutsche Industrie Normenausschuss	DIN
Deutsches Institute fur Normung	DIN
device under test	DUT
diameter	diam
dielectric absorption	DA
differential thermocouple voltmeter	DTVM
digital audio broadcasting	DAB
digital audio stationary head	DASH
digital audio tape	DAT
Digital Audio Video Council	DAVIC
digital audio workstations	DAW
digital compact cassette	DCC
digital data storage	DDS
digital home standard	DHS
digital light processing	DLP
digital micromirror device	DMD
digital phantom power	DPP
digital rights management	DRM
digital room correction	DRC
digital satellite system	DSS
digital signal processing	DSP
Digital Subscriber Line	DSL
digital sum value	DSV
digital to analog	D/A
digital TV	DTV
digital versatile disc	DVD
digital VHS	D-VHS
digital video	DV
digital video broadcasting	DVB
digital visual interface	DVI
digital voltmeter	DVM
digital-to-analog converter	DAC
direct broadcast	DBS
direct broadcast satellite	DBS
direct current	dc
direct current volts	V_{dc}
direct memory access	DMA
direct metal mastering	DMM
direct satellite broadcast	DSB
direct sound level in dB	dB_{DIR}
direct sound pressure level	L_D
direct stream digital	DSD
direct stream transfer	DST
direct time lock	DTL
direct to disk mastering	DDM
direct to home	DTH

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
directivity factor	Q
directivity index	DI
Discrete Fourier Transform	DFT
discrete multitone	DMT
display data channel	DDC
display power management signaling	DPMS
dissipation factor	DF
double sideband	DSB
dual expanded plastic insulated conductor	DEPIC
dual in-line package	DIP
dual-tone multifrequency	DTMF
dynamic host configuration protocol	DHCP
dynamic host control protocol	DHCP
dynamic noise reduction	DNR
dyne	dyn
EIA microphone sensitivity rating	G_M
eight-to-fourteen modulation	EFM
electrical metallic tubing	EMT
electrical power required	EPR
electrocardiograph	EKG
electromagnetic compatibility	ECM
electromagnetic interference	EMI
electromagnetic radiation	emr
electromagnetic unit	EMU
electromechanical relay	EMR
electromotive force	emf
electron volt	eV
electronic data processing	EDP
Electronic Field Production	EFP
Electronic Industries Alliance	EIA
Electronic Industries Association (obsolete)	EIA
electronic iris	E.I.
electronic music distribution	EMD
electronic news gathering	ENG
electronic voltohmmeter	EVOM
electronvolt	eV
electrostatic unit	ESU
Emergency Broadcast System	EBS
end of life vehicle	ELV
energy density level	L_W
energy frequency curve	EFC
energy level	L_E
energy-time-curve	ETC
Enhanced Definition Television	EDTV
enhanced direct time lock	DTLe
enhanced IDE	EIDE

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
environmental protection agency	EPA
equal level far end crosstalk	ELFEXT
equalizer	EQ
equipment distribution area	EDA
equivalent acoustic distance	EAD
equivalent input noise	EIN
equivalent rectangular bandwidth	ERB
equivalent resistance	R_{eq}
equivalent series inductance	ESL
equivalent series resistance	ESR
error checking and correcting random-access memory	ECC RAM
ethylene-propylene copolymer rubber	EPR
ethylene-propylene-diene monomer rubber	EPDM
European Broadcasting Union	EBU
expanded polyethylene-polyvinyl chloride	XPE-PVC
extended data out RAM	EDO RAM
extra-high voltage	EHV
extremely high frequency	EHF
extremely low frequency	ELF
far end crosstalk	FEXT
farad	F
Fast Discrete Fourier Transform	FDFT
Fast Fourier Transform	FFT
fast link pulses	FLPs
Federal Communications Commission	FCC
feedback stability margin	FSM
fiber data distributed interface	FDDI
fiber distribution frame	FDF
fiber optic connector	FOC
fiber optics	FO
fiber to the curb	FTTC
fiber to the home	FTTH
field programmable gate array	FPGA
field-effect transistor	FET
file transfer protocal	FTP
finite difference	FD
finite difference time domain	FDTD
finite impulse response	FIR
fire alarm and signal cable	FAS
Flame retardant ethylene propylene	FREP
flame retarded thermoplastic elastomer	FR-TPE
flexible OLED	FOLED
flexible organic light emitting diode	FOLED
Fluorinated ethylene-propylene	FEP
flux density	B

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
foot	ft/ [']
foot per minute	ft/min
foot per second	ft/s
foot per second squared	ft/s^2
foot pound-force	ft.lbf
foot poundal	ft.dl
footcandle	fc
footlambert	fL
forward error correction	FEC
four-pole, double-throw	4PDT
four-pole, single-throw	4PST
fractional part of	FRC
frame check sequence	FCS
frequency modulation	FM
frequency time curve	FTC
frequency-shift keying	FSK
frequency; force	f
frequently asked question	FAQ
full scale	FS
function indicator panel	FIP
gallon	gal
gallon per minute	gal/min
gauss	G
General Services Administration	GSA
gigacycle per second	GHz
gigaelectronvolt	GeV
gigahertz	GHz
gilbert	Gb
gram	g
Greenwich Mean Time	GMT
ground	GND
gypsum wallboard	GWB
head related transfer function	HRTF
Hearing Loss Association of America	HLAA
heating, ventilating, and air conditioning	HVAC
henry	H
hertz	Hz
high bit-rate digital subscriber line	HDSL
high definition - serial digital interface	HD-SDI
high definition multimedia interface	HDMI
high frequency	HF
high voltage	HV
high-bandwidth digital content protection	HDCP
high-definition multimedia interface	HDMI
high-definition television	HDTV
high-density linear converter system	HDLCS

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
high-pass filter	HPF
high-speed cable data service	HSCDS
high-speed parallel network technology	HIPPI
Home Automation and Networking Association	HANA
horizontal connection point	HCP
horizontal distribution areas	HDAs
horsepower	hp
hour	h
hybrid fiber/coaxial	HFC
hypertext markup language	HTML
hypertext transfer protocol	HTTP
ignition radiation suppression	IRS
impedance (magnitude)	Z
impulse response	IR
in the ear	ITE
inch	in, ["]
inch per second	in/s
independent consultants in audiovisual technology	ICAT
independent sideband	ISB
index matching gel	IMG
index of refraction	IOR
inductance	L
inductance-capacitance	LC
inductive reactance	X_L
inductor	L
infinite impulse response	IIR
infrared	IR
initial signal delay	ISD
initial time delay gap	ITDG
inner hair cells	IHC
input-output	I/O
inside diameter	ID
Institute of Electrical and Electronic Engineers	IEEE
Institute of Radio Engineers	IRE
instructional television fixed service	ITFS
Insulated Cable Engineers Association	ICEA
insulated gate field effect transistor	IGFET
insulated gate transistor	IGT
insulation displacement connector	IDC
insulation resistance	IR
integrated circuit	IC
integrated detectors/preamplifiers	IDP
integrated device electronics	IDE
integrated electronic component	IEC
integrated network management system	INMS
Integrated Services Digital Network	ISDN

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
intelligent power management system	IPM™
intensity level	L_I
interaural cross-correlation	IACC
interaural cross-correlation coefficient	IACC
interaural intensity difference	IID
interaural level difference	ILD
interaural phase difference	IPD
interaural time difference	ITD
intermediate frequency	IF
intermodulation	IM
intermodulation distortion	IM or IMD
international building code	IBC
International Communication Industries Association	ICIA
International Electrotechnical Commission	IEC
International Electrotechnical Engineers	IEEE
International Organization for Standardization	ISO
International Radio Consultative Committee	CCIR
International Standard Recording Code	ISRC
International Standards Organization	ISO
International Telecommunication Union	ITU
Internet Engineering Task Force	IETF
internet group management protocol	IGMP
internet protocol	IP
internet service provider	ISP
interrupted feedback (foldback)	IFB
inverse discrete Fourier transform	IDFT
IP Over Cable Data Network	IPCDN
ISDN digital subscriber line	IDSL
Japanese Standards Association	JSA
Joint Photographic Experts Group	JPEG
joule	J
joule per kelvin	J/K
junction field effect transistor	JFET
just noticeable difference	JND
kelvin	K
kilocycle per second	kHz
kiloelectronvolt	keV
kilogauss	kG
kilogram	kg
kilogram-force	kgf
kilohertz	kHz
kilohm	kΩ
kilojoule	kJ
kilometer	km
kilometer per hour	km/h

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
kilovar	kvar
kilovolt (1000 volts)	kV
kilovolt-ampere	kVA
kilowatt	kW
kilowatthour	kWh
knot	kn
lambert	L
large area systems	LAS
large-scale hybrid integration	LSHI
large-scale integration	LSI
lateral efficiency	LE
lateral fraction	LF
leadership in energy and environmental design	LEED
least significant bit	LSB
left, center, right, surrounds	LCRS
light amplification by stimulated emission of radiation	LASER
light dependent resistor	LDR
light emitting diode	LED
linear time invariant	LTI
liquid crystal display	LCD
liquid crystal on silicon	LCoS
listening environment diagnostic recording	LEDR
liter	l
liter per second	l/s
live end–dead end	LEDE
local area network	LAN
local exchange carrier	LEC
local multipoint distribution service	LMDS
logarithm	log
logarithm, natural	ln
long play	LP
look up table	LUT
loudspeaker sensitivity	L_{sens}
low frequency	LF
low frequency effects	LFE
low power radio services	LPRS
low-frequency effects	LFE
low-pass filter	LPF
lower sideband	LSB
lumen	lm
lumen per square foot	lm/ft^2
lumen per square meter	lm/m^2
lumen per watt	lm/W
lumen second	lm.s
lux	lx

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
magneto hydrodynamics	MHD
magneto-optical recording	MOR
magneto-optics	MO
magnetomotive force	MMF
mail transfer protocol	MTP
main cross connect	MC
main distribution areas	MDA
manufacturing automation protocol	MAP
Mass Media Bureau	MMB
master antenna television	MATV
matched resistance	R_M
maxwell	Mx
mean free path	MFP
media access control	MAC
medium area systems	MAS
medium frequency	MF
megabits per second	Mbps
megabyte	MB
megacycle per second	MHz
megaelectronvolt	MeV
megahertz	MHz
megavolt	MV
megawatt	MW
megohm	MΩ
metal-oxide semiconductor	MOS
metal-oxide semiconductor field-effect transistor	MOSFET
metal-oxide varistor	MOV
meter	m
meter-kilogram-second	MKS
metropolitan area network	MAN
microampere	µA
microbar	µbar
microelectromechanical systems	MEMS
microfarad	µF
microgram	µg
microhenry	µH
micrometer	µm
micromho	µmho
microphone	mic
microsecond	µs
microsiemens	µS
microvolt	µV
microwatt	µW
midi time code	MTC
mile (statute)	mi
mile per hour	mi/h
milli	m
milliampere	mA
millibar	mbar
millibarn	mb
milligal	mGal
milligram	mg
millihenry	mH
milliliter	ml
millimeter	mm
millimeter of mercury, conventional	mmHg
millisecond	ms
millisiemens	mS
millivolt	mV
milliwatt	mW
minidisc	MD
minute (plane angle)	...′
minute (time)	min
modified rhyme test	MRT
modulation reduction factor	m(F)
modulation transfer function	MTF
modulation transmission function	MTF
mole	mol
most significant bit	MSB
motion drive amplifier	MDA
motion JPEG	M-JPEG
Motion Picture Experts Group	MPEG
moves, adds, and changes	MACs
moving coil	MC
multichannel audio digital interface	MADI
multichannel reverberation	MCR
multichannel audio digital interface	MADI
Multimedia Cable Network System Partners Ltd	MCNS
multiple system operator	MSO
multiple-in/multiple-out	MIMO
multiplier/accumulator	MAC
multipoint distribution system	MDS
multistage noise shaping	MASH
multiuser telecommunications outlet assembly	MUTOA
music cd plus graphics	CD-G
musical instrument digital interface	MIDI
mutual inductance	M
nanoampere	nA
nanofarad	nF
nanometer	nm
nanosecond	ns
nanowatt	nW

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
National Association of Broadcasters	NAB
National Association of the Deaf	NAD
National Broadcasting Company	NBC
National Bureau of Standards	NBS
National Electrical Code	NEC
National Electrical Contractors Association	NECA
National Electrical Manufacturers Association	NEMA
National Fire Protection Association	NFPA
National Institute Of Occupational Safety And Health	NIOSH
National Systems Contractors Association	NSCA
National Television Standards Committee	NTSC
National Television System Committee	NTSC
near end cross talk	NEXT
near-instantaneous companding	NICAM
needed acoustic gain	NAG
negative-positive-negative	NPN
neper	Np
network address translation	NAT
network operations center	NOC
neutral density filter	N/D
newton	N
newton meter	N·m
newton per square meter	N/m^2
no-epoxy/no-polish	NENP
noise figure; noise frequency	NF
noise reduction coefficient	NRC
noise voltage	E_n
noise-operated automatic level adjuster	NOALA
nonreturn-to-zero inverted	NRZI
normal link pulses	NLPs
number of open microphone	NOM
numerical aperture	NA
oersted	Oe
Office de Radiodiffusion Television Francois	ORTF
ohm	Ω
on-screen manager	OSM™
open system interconnect	OSI
open-circuit voltage	E_o
operational transconductance amplifier	OTA
operations support systems	OSS
opposed current interleaved amplifier	OCIA
optical carrier	OC
optical time domain reflectometer	OTDR
optimized common mode rejection	OCMR
optimum power calibration	OPC

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
optimum source impedance	OSI
optoelectronic integrated circuit	OEIC
organic light emitting diode	OLED
orthogonal frequency division multiplexing	OFDM
ounce (avoirdupois)	oz
outer hair cells	OHC
output level in dB	L_{out}
output voltage	E_{OUT}
outside diameter	OD
oxygen-free, high-conductivity copper	OFHC
pan/tilt/zoom	PTZ
parametric room impulse response	PRIR
pascal	Pa
peak amplitude	A_p
peak program meter	PPM
peak reverse voltage	PRV
peak-inverse-voltage	piv
peak-reverse-voltage	prv
peak-to-peak amplitude	A_{p-p}
percentage of articulation loss for consonants	%Alcons
Perfluoroalkoxy	PFA
permanent threshold shift	PTS
personal listening systems	PLS
phase alternation line	PAL
phase angle	I
phase frequency curve	PFC
phase locked loop	PLL
phase modulation	PM
phonemically balanced	PB
physical medium dependent	PMD
pickup	PU
picoampere	pA
picofarad	pF
picosecond	ps
picowatt	pW
picture in picture	PIP
pinna acoustic response	PAR
pint	pt
plain old telephone service	POTS
plasma	PDP
plastic insulated conductor	PIC
plate current	I_p
plate efficiency	E_{ff}
plate resistance	r_p
plate voltage	E_p
point-to-point protocol	PPP

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
polarization beam splitter	PBS
polyethylene	PE
polyethylene aluminum steel polyethylene	PASP
Polypropylene	PP
Polyurethane	PUR
polyvinyl chloride	PVC
polyvinylidene fluoride	PVDF
positive-negative-positive	PNP
positive, intrinsic, negative	PIN
potential acoustic gain	PAG
pound	lb
pound (force) per square inch. Although the use of the abbreviation psi is common, it is not recommended.	lbf/in^2, psi,
pound-force	lbf
pound-force foot	lb·fft
poundal	pdl
power backoff	PBO
power calibration area	PCA
power factor	PF
power factor correction	PFC
power level	L_W, dB-PWL
power out	P_o
power over ethernet	PoE
power sourcing equipment	PSE
power sum alien ELFEXT	PSAELFEXT
power sum alien equal level far-end crosstalk	PSAELFEXT
Power sum alien far-end crosstalk	PSAFEXT
Power sum alien near-end crosstalk	PSANEXT
power sum alien NEXT	PSANEXT
powered devices	PD
preamplifier	preamp
precision adaptive subband coding	PASC
precision audio link	PAL
prefade listen	PFL
primary rate interface ISDN	PRI
printed circuit	PC
private branch exchange	PBX
Professional Education and Training Committee	PETC
programmable gate array	PGA
programmable logic device	PLD
programmable read-only memory	PROM
public switched telephone network	PSTN
pulse code modulation	PCM
pulse density modulation	PDM
pulse end modulation	PEM
pulse-amplitude modulation	PAM
pulse-duration modulation	PDM
pulse-frequency-modulation	PFM
pulse-position modulation	PPM
pulse-repetition frequency	PRF
pulse-repetition rate	PRR
pulse-time modulation	PTM
pulse-width modulation	PWM
quadratic residue diffuser	QRD
quality factor	Q
quality of service	QoS
quandrature amplitude modulation	QAM
quart	qt
quarter wave plate	QWP
quaternary phase shift keying	QPSK
rad	rd
radian	rad
radio data service	RDS
radio frequency	RF
radio frequency identification	RFID
radio information for motorists	ARI
radio-frequency interference	RFI
rambus DRAM	Rambus, RDRAM
random access memory	RAM
random-noise generator	RNG
rapid speech transmission index	RASTI
reactance	X
read only memory	ROM
real-time analyzer	RTA
real-time transport protocol	RTP
Recording Industry Association of America	RIAA
recording management area	RMA
red, green, blue	RGB
redundant array of independent disks	RAID
reflection-free zone	RFZ
reflections per second	RPS
Regional Data Center	RDC
registered communication distribution designer	RCDD
remote authentication dial-in user service	RADIUS
report on comments	ROC
report on proposal	ROP
request for proposals	RFP
resistance-capacitance	RC
resistance-inductance-capacitance	RLC
resistor	R

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
resistor-capacitor	RC
resistor-transistor logic	RTL
resource reservation protocol	RSVP
restriction of hazardous substances	RoHS
return loss	RL
reverberant sound level in dB	L_R
reverberation time	RT_{60}
revolution per minute	r/min, rpm
revolution per second	r/s, rps
ripple factor	Y
robust service network	RSN
roentgen	R
room constant	Sa
root-mean-square	rms
root-mean-square voltage	Vrms
rotary digital audio tape	R-DAT
rotary head digital audio tape	R-DAT
round conductor flat Cable	RCFC
sample and hold	S/H
sample-rate convertor	SRC
satellite news gathering	SNG
Screen Actors Guild	SAG
screened twisted pair	ScTP
second (plane angle)	″
second (time)	s
second audio program	SAP
secure digital music initiative	SDMI
self-monitoring analysis and reporting technology	SMART
sensitivity	sensi
Sequence Electronique Couleur Avec Memoire	SECAM
serial copy management system	SCMS
serial digital interface	SDI
serial digital video	SDV
service station identifier	SSID
shield current induced noise	SCIN
shielded twisted pair(s)	STP
short noise	i_{sn}
short wave	SW
siemens	S
signal delay	SD
signal-to-noise ratio	SNR
silicon controlled rectifier	SCR
simple control protocol	SCP
simple network management protocol	SNMP
single in-line package	SIP
single sideband	SSB

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
single-pair high bit-rate digital subscriber line	S-HDSL
single-pole, double-throw	SPDT
single-pole, single-throw	SPST
small computer system	SCSI
Society of Automotive Engineers	SAE
Society of Motion Picture & Television Engineers	SMPTE
solid state music	SSM
solid state relay	SSR
song position pointer	SPP
sound absorption average	SAA
sound level meter	SLM
sound pressure in dB	$\mathrm{dB}_{\mathrm{SPL}}$
sound pressure level	L_{p}, SPL
sound transmission class	STC
sound, audiovisual, and video integrators	SAVVI
source resistance	R_{s}
speech transmission index	STI
square foot	ft^2
square inch	in^2
square meter	m^2
square yard	yd^2
standard definition - serial digital interface	SD-SDI
standard definition television	SDTV
standing-wave ratio	SWR
static contact resistance	SCR
static RAM	SRAM
steradian	sr
storage area network	SAN
structural return loss	SRL
structured cabling system	SCS
stubs wire gage	SWG
subminiature A connector	SMA
subminiature B connector	SMB
subminiature C connector	SMC
Subsidiary Communications Authorization	SCA
super audio CD	SACD
super audio compact disc	SACD
super video home system	Super VHS
super-luminescent diode	SLD
super-high frequency	SHF
symmetric digital subscriber line	SDSL
synchronous code division multiple access	S-CDMA
synchronous optical network	SONET
table of contents	TOC
telecommunications enclosure	TE
Telecommunications Industry Association	TIA

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
telecommunications room	TR
television	TV
television interference	TVI
temperature differential	ΔT
temporary threshold shift	TTS
tesla	T
tetrafluoroethylene	TFE
thermal noise	TN, i_{tn}
thermocouple; time constant	TC
thin film transistors	TFT
thousand circular mils	kcmil
three-pole, double-throw	3PDT
three-pole, single-throw	3PST
time	T
time delay spectrometry	TDS
time division multiple access	TDMA
time division multiplexing	TDM
time energy frequency	TEF
timebase corrector	TBC
ton	ton
tonne	t
total harmonic distortion	THD
total harmonic distortion plus noise	THD+N
total sound level in dB	L_T
total surface area	S
transient intermodulation distortion	TIM
transistor-transistor logic	TTL
transmission control protocol/internet protocol	TCP/IP
transmission loss	TL
transparent OLED	TOLED
transparent organic light Emitting diode	TOLED
transverse electric	TE
transverse electromagnetic	TEM
transverse magnetic	TM
traveling-wave tube	TWT
TV receive only	TVRO
twisted pair-physical medium dependent	TP-PMD
ultrahigh frequency	UHF
ultraviolet	UV
Underwriters Laboratories, Inc.	UL
uniform building code	UBC
unit interval	UI
unit of absorption	Sabin
universal disc format	UDF
Universal Powerline Association	UPA
universal serial bus	USB

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
universal service order code	USOC
unshielded twisted pair(s)	UTP
upper sideband	USB
user datagram protocol	UDP
user defined protocol	UDP
vacuum-tube voltmeter	VTVM
variable constellation/multitone modulation	VC/MTM
variable speed oscillator	VSO
variable-frequency oscillator	VFO
velocity of propagation	VP
vertical-cavity surface-emitting laser	VCSEL
very high bit rate digital subscriber line	VDSL
very high frequency	VHF
very low frequency	VLF
vibratory acceleration level	L_a
vibratory force level	L_F
vibratory velocity level	L_v
Video Electronics Standards Association	VESA
video graphics array	VGA
video home system	VHS
video on demand	VOD
video RAM	VRAM
virtual local area network	VLAN
virtual private networks	VPN
voice over internet protocol	VoIP
voice over wireless fidelity	VoWi-Fi
volt	V
volt-ohm-milliammeter	VOM
voltage (electromotive force)	E
voltage controlled crystal oscillator	VCXO
voltage gain	μ
voltage standing wave ratio	VSWR
voltage-controlled amplifier	VCA
voltage-controlled oscillator	VCO
voltampere	VA
volume indicator	VI
volume unit	VU
watt	W
watt per steradian	W/sr
watt per steradian square meter	W/(sr·m^2)
watthour	Wh
wavelength	M
wavelength division multiplexing	WDM
weber	Wb
weighted modulation transmission function	WMTF
wide area network	WAN

Table 48-9. Recommended Abbreviations

Unit or Term	Symbol or Abbreviation
windows media audio	WMA
wired equivalent privacy	WEP
wireless access points	WAPs
wireless application protocol	WAP
wireless communications service	WCS
wireless fidelity	WiFi
wireless microwave access	WiMax
World Health Organization	WHO
write once	WO
write once read many	WORM
yard	yd
zone distribution area	ZDA

48.12 Audio Frequency Range

The audio spectrum is usually considered the frequency range between 20 Hz and 20 kHz, Fig. 48-9. In reality, the upper limit of hearing pure tones is between 12 kHz and 18 kHz, depending on the person's age and sex and how well the ears have been protected against loud sounds. Frequencies above 20 kHz cannot be heard as a sound, but the effect created by such frequencies (i.e., rapid rise time) can be heard.

48.13 Surface Area and Volume Equations

To find the *surface area and volume* of complex areas, the area can often be divided into a series of simpler areas and handled one at a time. Figs. 48-10A–H are equations for various and unusual volumes.

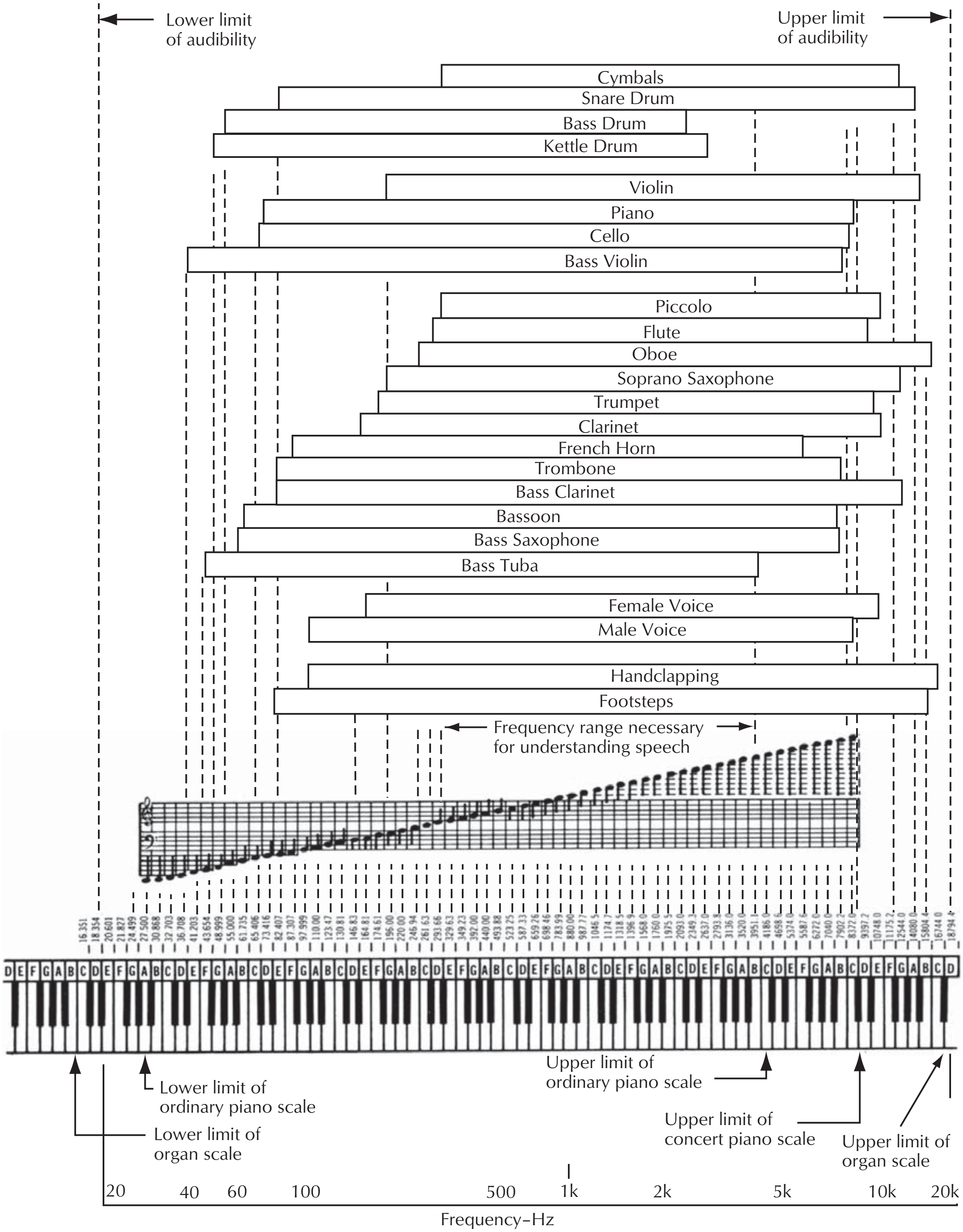


Figure 48-9. Audible frequency range.

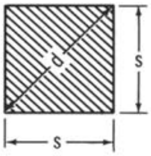
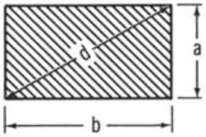
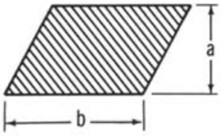
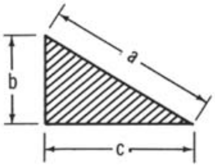
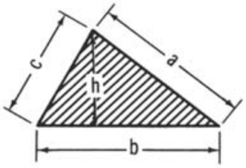
 Square	A = area. $A = s^2$ $A = \frac{1}{2} d^2$ $s = 0.7071\ d = \sqrt{A}$ $d = 1.414\ s = 1.414 \sqrt{A}$
 Rectangle	A = area. $A = ab$ $A = a \sqrt{d^2-a^2} = b \sqrt{d^2-b^2}$ $d = \sqrt{a^2+b^2}$ $a = \sqrt{d^2-b^2} = A \div b$ $b = \sqrt{d^2-a^2} = A \div a$
 Parallelogram	A = area. $A = ab$ $a = A \div b$ $b = A \div a$ Note that dimension a is measured at right angles to line b.
 Right-angled triangle	A = area. $A = \frac{bc}{2}$ $a = \sqrt{b^2+c^2}$ $b = \sqrt{a^2-c^2}$ $c = \sqrt{a^2-b^2}$
 Acute-angled triangle	A = area. $A = \frac{bh}{2} = \frac{b}{2} \sqrt{a^2-\left(\frac{a^2+b^2-c^2}{2b}\right)^2}$ If $S = \frac{1}{2}(a+b+c)$, then $A = \sqrt{S(S-a)(S-b)(S-c)}$

Figure 48-10. Equations for finding surface areas for complex shapes.

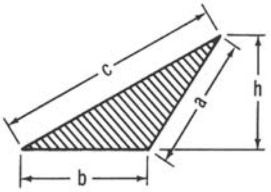
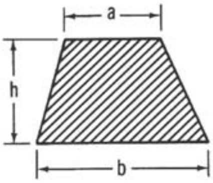
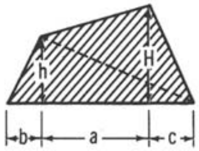
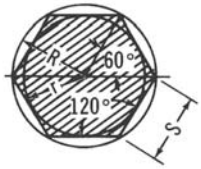
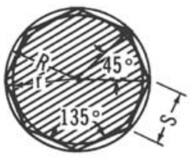
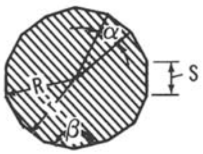
 Obtuse-angled triangle	A = area. $A = \frac{bh}{2} = \frac{b}{2}\sqrt{a^2 - \left(\frac{c^2-a^2-b^2}{2b}\right)^2}$ If $S = \frac{1}{2}(a+b+c)$, then $A = \sqrt{S(S-a)(S-b)(S-c)}$
 Trapezoid	A = area. $A = \frac{(a+b)h}{2}$
 Trapezium	A = area. $A = \frac{(H+h)a+bh+cH}{2}$ A trapezium can also be divided into two triangles as indicated by the dotted line. The area of each of these triangles is computed, and the results added to find the area of the trapezium.
 Regular hexagon	A = area; R = radius of circumscribed circle; r = radius of inscribed circle. $A = 2.598\ s^2 = 2.598\ R^2 = 3.464\ r^2$ $R = s = 1.1155r$ $r = 0.866\ s = 0.866\ R$ $s = R = 1.1155\ r$
 Regular octagon	A = area; R = radius of circumscribed circle; r = radius of inscribed circle; $A = 4.828\ s^2 = 2.828\ R^2 = 3.314\ r^2$ $R = 1.307\ s = 1.082\ r$ $r = 1.207\ s = 0.924\ R$ $s = 0.765\ R = 0.828\ r$
 Regular polygon	A = area; n = number of sides. $\alpha = 360° \div n$ $\beta = 180° - \alpha$ $A = \frac{nsr}{2} = \frac{ns}{2}\sqrt{R^2 - \frac{s^2}{4}}$ $R = \sqrt{r^2 + \frac{s^2}{4}}$; $r = \sqrt{R^2 - \frac{s^2}{4}}$; $s = 2\sqrt{R^2 - r^2}$

Figure 48-11. Equations for finding surface areas for complex shapes.

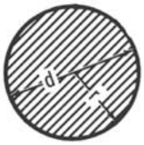
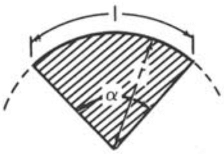
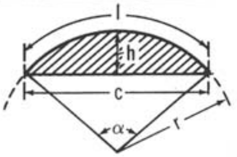
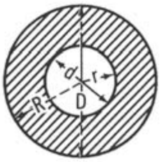
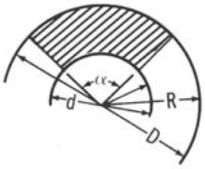
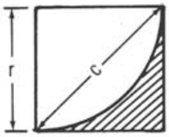
 Circle	A = area; C = circumference. $A = \pi r^2 = 3.1416r^2 = 0.7854\ d^2$ $C = 2\pi r = 6.2832\ r = 3.1416\ d$ $r = C \div 6.2832 = \sqrt{A \div 3.1416} = 0.564\ \sqrt{A}$ $d = C \div 3.1416 = \sqrt{A \div 0.7854} = 1.128\ \sqrt{A}$ Length of arc for center-angle of 1° = 0.008727 d Length of arc for center-angle of n° = 0.008727 nd
 Circular sector	A = area; l = length of arc; α = angle, in degrees. $l = \frac{r \times \alpha \times 3.1416}{180} = 0.01745\ r\alpha = \frac{2A}{r}$ $A = \frac{1}{2}\ rl = 0.008727\ \alpha r^2$ $\alpha = \frac{57.296\ l}{r} \qquad r = \frac{2A}{l} = \frac{56.296\ l}{\alpha}$
 Circular segment	A = area; l = length of arc; α = angle, in degrees. $c = 2\ \sqrt{h(2r-h)} \qquad A = \frac{1}{2}\ [rl - c(r-h)]$ $r = \frac{c^2 + 4\ h^2}{8h} \qquad l = 0.01745\ r\alpha$ $h = r - \frac{1}{2}\ \sqrt{4r^2 - c^2} \qquad \alpha = \frac{57.296l}{r}$
 Circular ring	A = area. $A = \pi\ (R^2 - r^2) = 3.1416\ \ (R^2 - r^2)$ $= 3.1416\ R + r)\ (R - r)$ $= 0.7854\ (D^2 - d^2) = 0.7854\ (D + d)\ (D - d)$
 Circular ring sector	A = area; α = angle, in degrees. $A = \frac{\alpha\pi}{360}\ (R^2 - r^2) = 0.00873\alpha\ (R^2 - r^2)$ $= \frac{\alpha\pi}{4 \times 360}\ (D^2 - d^2) = 0.00218\ \alpha\ (D^2 - d^2)$
 Spandrel or fillet	A = area. $A = r^2 - \frac{\pi r^2}{4} = 0.215\ r^2$ $= 0.1075\ c^2$

Figure 48-12. Equations for finding surface areas for complex shapes.

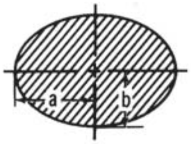
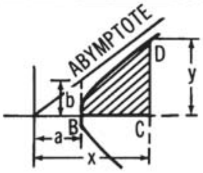
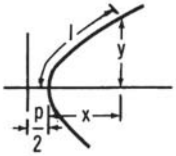
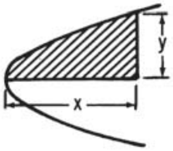
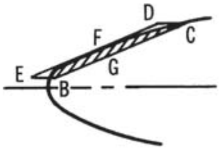
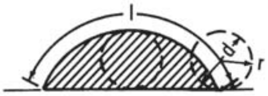
 Elipse	A = area; P = perimeter or circumference. $A = \pi ab = 3.1416\ ab.$ An approximate formula for the perimeter is: $P = 3.1416\sqrt{2(a^2+b^2)}$
 Hyperbola	A = area BCD $A = \frac{xy}{2} - \frac{ab}{2}\ \text{hyp. log}\left(\frac{x}{a} + \frac{y}{b}\right)$
 Parabola	l = length of arc. $l = \frac{p}{2}\left[\sqrt{\frac{2x}{p}\left(I + \frac{2x}{p}\right)} + \text{hyp. log}\sqrt{\frac{2x}{p} + I + \frac{2x}{p}}\right)\Big]$ When x is small in proportion to y, the following is a close approximation: $l = y\left[I \div \frac{2}{3}\left(\frac{x}{y}\right)^2 - \frac{2}{5}\left(\frac{x}{y}\right)^4\right]$ or $l = \sqrt{y^2 + \frac{4}{3}x^2}$
 Parabola	A = area. $A = \frac{2}{3}\ xy$ (The area is equal to two-thirds of the rectangle which has x for its base and y for its height.)
 Segment of Parabola	A = area. Area BFC = A = ⅔ area of parallelogram BCDE. If FG is the height of the segment, measured at right angles to BC, then: Area of segment BFC = ⅔ BC×FG
 Cycloid	A = area; l = length of cycloid. $A = 3\pi r^2 = 9.4248\ r^2 = 2.3562\ d^2$ = 3 × area of generating circle $l = 8\ r = 4\ d$

Figure 48-13. Equations for finding surface areas for complex shapes.

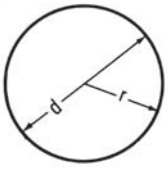
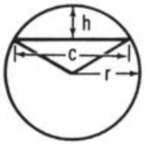
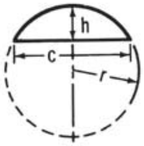
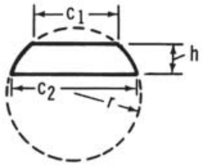
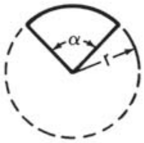
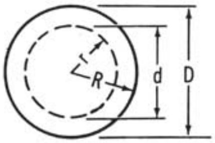
 Sphere	V = volume; A = area of surface. $V = \frac{4\pi r^2}{3} = \frac{\pi d^3}{6} = 4.1888\, r^3 = 0.5236\, d^3$ $A = 4\pi r^2 = \pi d^2 = 12.5664 r^2 = 3.1416\, d^2$ $r = \sqrt[3]{\frac{3V}{4\pi}} = 0.6204 \sqrt{V}$
 Spherical sector	V = volume; A = total area of conical and spherical surface. $V = \frac{2\pi r^2 h}{3} = 2.0944 r^2 h$ $A = 3.1416 r\, (2h + \frac{1}{2}\, c)$ $c = 2\sqrt{h\,(2r-h)}$
 Spherical segment	V = volume; A = area of spherical surface. $V = 3.1416\, h^2 \left(r - \frac{h}{3}\right) = 3.1416 h \left(\frac{c^2}{8} + \frac{h^2}{6}\right)$ $A = 2\pi rh = 6.2832 rh = 3.1416 \left(\frac{c^2}{4} + h^2\right)$ $c = 2\sqrt{h(2r-h)};\ r = \frac{c^2 + 4h^2}{8h}$
 Spherical zone	V = volume; A = area of spherical surface. $V = 0.5236\, h \left(\frac{3{c_1}^2}{4} + \frac{3{c_2}^2}{4} + h^2\right)$ $A = 2\pi rh = 62832\, rh$ $r = \sqrt{\frac{{c_2}^2}{4} + \left(\frac{{c_2}^2 - {c_1}^2 - 4h^2}{8h}\right)^2}$
 Spherical wedge	V = volume; A = area of spherical surface; α = center angle in degrees. $V = \frac{\alpha}{360} \times \frac{4\pi r^3}{3} = 0.0116\, \alpha r^3$ $A = \frac{\alpha}{360} \times 4\pi r^2 = 0.0349\, \alpha r^2$
 Hollow sphere	V = volume. $V = \frac{4\pi}{3}\, (R^3 - r^3) = 4.1888\, (R^3 r^3)$ $= \pi\ 6\ \ (D^3 - d^3) = 0.5236\, (D^3 - d^3)$

Figure 48-14. Equations for finding surface areas for complex shapes.

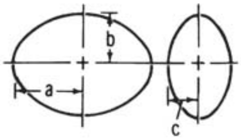
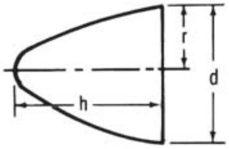
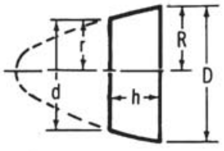
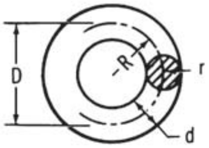
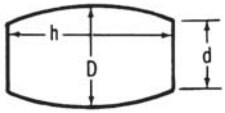
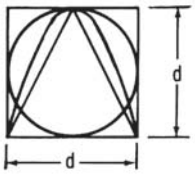
 Ellipsoid	V = volume; A = area of surface. $V = \frac{4\pi}{3}$ abc = 4.1888 abc In an ellipsoid of revolution, or spheroid, where $b = c$: $V = 4.1888\ ab^2$, and $A = \frac{4\pi}{\sqrt{2}}\ b\ \sqrt{a^2+b^2}$
 Paraboloid	V = volume; $V = \frac{1}{2}\ \pi r^2 h = 0.3927\ d^2 h$ A = area; $A = \frac{2\pi}{3p}\left[\sqrt{\left(\frac{d^2}{4} + p^2\right)^3} - p^3\right]$ in which $p = \frac{d^2}{8h}$
 Parabolidal segment	V = volume. $V = \frac{\pi}{2}\ h\ (R^2+r^2) = 1.5708\ h\ (R^2+r^2)$ $= \frac{\pi}{8}\ h\ (D^2+d^2) = 0.3927\ h\ (D^2+d^2)$
 Torus	V = volume; A = area of surface. $V = 2\pi^2\ Rr^2 = 19.739\ Rr^2$ $= \frac{\pi^2}{4}\ Dd^2 = 2.4674\ Dd^2$ $A = 4\pi^2\ Rr = 39.478\ Rr$ $= \pi^2 Dd = 9.8696\ Dd$
 Barrel	V = approximate volume. If the sides are bent to the arc of a circle: $V = \frac{1}{12}\ \pi h\ (2\ D^2+d^2) = 0.262\ h\ (2\ D^2+d^2)$ If the sides are bent to the arc of a parabola: $V = 0.209\ h\ (2\ D^2+Dd+\frac{3}{4}\ d^2)$
	If d = base diameter and height of a cone, a paraboloid and a cylinder, and the diameter of a sphere, then the volumes of these bodies are to each other as below: Cone: paraboloid: sphere: cylinder = $\frac{1}{3} : \frac{1}{2} : \frac{2}{3} : 1$

Figure 48-15. Equations for finding surface areas for complex shapes.

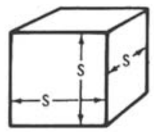
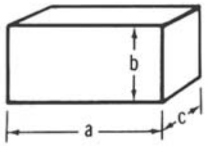
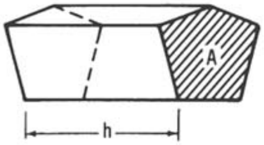
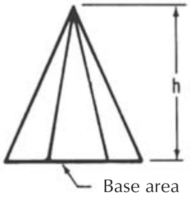
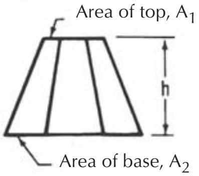
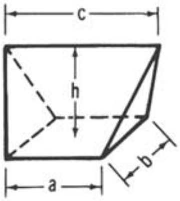
 Cube	V = volume. $V = s^3$ $s = \sqrt[3]{V}$
 Square prism	V = volume. $V = abc$ $a = \frac{V}{bc}$ $b = \frac{V}{ac}$ $c = \frac{V}{ab}$
 Prism	V = volume; A = area of end surface. $V = h \times A$ The area A of the end surface is found by the formulas for areas of plane figures on the preceding pages. Height h must be measured perpendicular to end surface.
 Pyramid	V = volume $V = \frac{1}{3} h \times \text{area of base.}$ If the base is a regular polygon with n sides, and s = length of side, r = radius of inscribed circle, and R = radius of circumscribed circle, then: $V = \frac{nsrh}{6} = \frac{nsh}{6}\sqrt{R^2 - \frac{s^2}{4}}$
 Frustum of pyramid	V = volume. $V = \frac{h}{3}(A_1 + A_2 + \sqrt{A_1 \times A_2})$
 Wedge	V = volume. $V = \frac{(2a+c)bh}{6}$

Figure 48-16. Equations for finding surface areas for complex shapes.

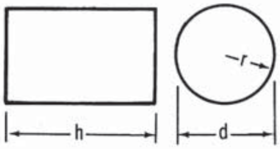
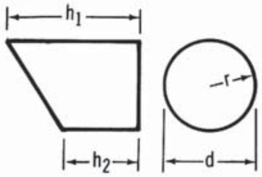
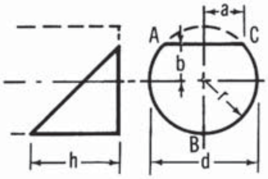
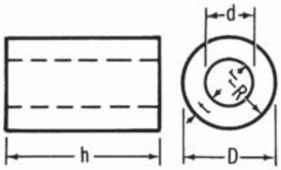
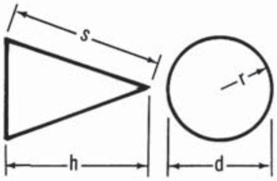
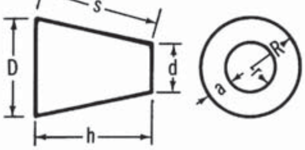
 Cylinder	V = volume; S = area of cylindrical surface. $V = 3.1416\ r^2h = 0.7854\ d^2h$ $S = 6.2832\ rh = 3.1416\ dh$ Total area A of cylindrical surface and end surfaces: $A = 6.2832\ r(r+h) = 3.1416\ d\left(\frac{1}{2}d+h\right)$
 Portion of cylinder	V = volume; S = area of cylindrical surface. $V = 1.5708\ r^2(h_1+h_2) = 0.3927\ d^2\ (h_1+h_2)$ $S = 3.1416\ r\ (h_1+h_2) = 1.5708\ d(h_1+h_2)$
 Portion of cylinder	V = volume; S = area of cylindrical surface. $V = \left(\frac{2}{3}a^3 \pm b \times \text{area ABC}\right)\frac{h}{r \pm b}$ $S = (ad \pm b \times \text{length of arc ABC})\ \frac{h}{r \pm b}$ Use + when base area is larger, and − when base area is less than one-half the base circle.
 Hollow cylinder	V = volume. $V = 3.1416\ h(R^2-r^2) = 0.7854\ h(D^2-d^2)$ $= 3.1416\ ht\ (2R-t) = 3.1416\ ht(D-t)$ $= 3.1416\ ht\ (2r+t) = 3.1416\ ht(d+t)$ $= 3.1416\ ht\ (R+r) = 1.5708\ ht(D+d)$
 Cone	V = volume; A = area of conical surface $V = \frac{3.1416\ r^2h}{3} = 1.0472\ r^2h = 0.2618\ d^2h$ $A = 3.1416\ r\sqrt{r^2+h^2} = 3.1416rs = 1.5708\ ds$ $s = \sqrt{r^2+h^2} = \sqrt{\frac{d^2}{4}+h^2}$
 Frustum of cone	V = volume; A = area of conical surface. $V = 1.0472\ h(R^2+Rr+r^2) = 0.2618\ h(D^2+Dd+d^2)$ $A = 3.1416\ s(R+r) = 1.5708\ s(D+d)$ $a = R-r \qquad s = \sqrt{a^2+h^2} = \sqrt{(R-r)^2+h^2}$

Figure 48-17. Equations for finding surface areas for complex shapes.

Bibliography

Metric Units of Measure and Style Guide, U.S. Metric Association, Inc., Sugarloaf Star Route, Boulder, CO 80302.

V. Antoine, *Guidance for Using the Metric System* (SI Version), Society for Technical Communication, 1010 Vermont Avenue, NW, Washington, DC 20005.

D. Davis, "Synergetic Audio Concepts," vol. 5, no. 2, October 1977.

John L. Feirer, *SI Metric Handbook*, The Metric Company.

Handbook of Electronics Tables and Formulas, 6th Edition, Carmel IN: SAMS, a division of Macmillan Computer Publishing, 1986.

D. Bohn, "The Bewildering Wilderness," *S&VC,* September 2000.

Index

Symbols

294
%Alcons 1391–1392, 1398, 1400, 1404, 1408
  acoustics for auditoriums and concert halls 157, 159
  designing for speech intelligibility 1391, 1398
  intelligibility
    Alcons 1396
  loudspeaker clusters 654
  Test and measurement 1625
  units of measurement 1659

Numerics

1/F noise
  consoles 861
1000Base-LH
  digital audio interfacing and networks 1490
1000Base-LX
  digital audio interfacing and networks 1490
1000Base-SX
  digital audio interfacing and networks 1490
1000Base-T
  digital audio interfacing and networks 1490
100BaseT 1490
  digital audio interfacing and networks 1490
100Base-TX
  digital audio interfacing and networks 1490
10Base2 1489
  digital audio interfacing and networks 1489
10Base5 1489
  digital audio interfacing and networks 1489
10BaseF
  digital audio interfacing and networks 1490
10BaseT 1489, 1501
  digital audio interfacing and networks 1489
10Broad36 1489
  digital audio interfacing and networks 1489
16 track console routing 854
1Base2 1489
  digital audio interfacing and networks 1489
32 track routing
  consoles 854
5.1 surround sound
  surround sound 1593
7.1 surround sound format
  surroound sound 1598
70.7 V loudspeaker transformer
  sound system design 1291
70.7 V or 100 V loudspeaker system
  sound system design 1291
70.7 volt loudspeaker system
  wire and cable 435

A

A/D conversion
  consoles 942, 954
  gain structure 1230
AACS
  optical disc formats 1152
Abbreviations
  units of measurement 1667
Absolute levels
  fundamentals of audio and acoustics 25
Absolute phase
  consoles 891
Absolute pitch
  psychoacoustics 56
Absorbed glass mat batteries (AGM) 696
Absorber data
  computer aided sound system design 1354
Absorber sound power
  acoustics for auditoriums and concert halls 151
Absorption
  acoustical modeling and auralization 219
  acoustical treatment for indoor areas 97
  fiber optics 459
  from trees and shrubs 230
  fundamentals of audio and acoustics 29
  gypsum cavity wall absorption 229
  Sabine 135
Absorption coefficient 222, 229
  acoustical modeling and auralization 217, 219
  acoustical noise control 93
  designing for speech intelligibility 1396
  units of measurement 1660
Absorption coefficients of building materials and absorbers
  acoustical treatment for indoor areas 115
Absorption of sound by air 205
Absorption ratings
  acoustical treatment for indoor areas 101
Absorption testing
  acoustical treatment for indoor areas 98
ac circuits
  grounding and interfacing 1182–1183
Ac current in various countries
  wire and cable 445
Ac differential relay 381
Ac power cords and receptacles
  wire and cable 443
Ac receptacles
  sound system design 1302
Ac relays 380
Ac resistance
  of resistors 245
Ac ripple current rating
  of a capacitor 258
Acceleration

  units of measurement 1652
Acoustic boundaries
  loudspeakers 625
Acoustic Control System (ACS)
  acoustics for auditoriums and concert halls 193
Acoustic delay
  delay 809
Acoustic feedback
  analog discs 1035
Acoustic gain
  fundamentals of audio and acoustics 32
  sound system design 1240
Acoustic impedance
  fundamentals of audio and acoustics 28
Acoustic lens
  loudspeakers 614
Acoustic power
  loudspeakers 633
Acoustic pressure levels
  fundamentals of audio and acoustics 26
Acoustic programs
  EASE 648, 651
  Noise Level Analysis 70
  SMAART 1309, 1334
  TEF 20 67
  TEF 25 70
acoustic programs
  SMAART 1618
Acoustic reciprocity
  acoustical modeling and auralization 226
Acoustic reflections
  fundamentals of audio and acoustics 28
Acoustic reflex
  psychoacoustics 45
Acoustic sources
  fundamentals of audio and acoustics 36
Acoustic wave
  fundamentals of audio and acoustics 28
Acoustic wavelengths
  fundamentals of audio and acoustics 30–31
  test and measurement 1619
Acoustical absorption
  acoustical treatment for indoor areas 97
Acoustical baffles
  acoustical treatment for indoor areas 108
Acoustical banners
  acoustical treatment for indoor areas 109
Acoustical crosstalk
  interpretation and tour group systems 1534
Acoustical diffusion
  acoustical treatment for indoor areas 115
Acoustical doors
  acoustical noise control 83
Acoustical modeling
  acoustical modeling and auralization 215
Acoustical modeling and auralization
  absorption 219
  absorption coefficient 217, 219
  absorption from trees and shrubs 230
  acoustic reciprocity 226
  angle of incidence 222
  auralization 215, 231
  beam-tracing models 224
  binaural impulse response (BIR) 232
  binaural reproduction 232
  binaural reproduction using headphones 232
  boundary element analysis 227
  computational models 215, 220
  damping materials 228
  diffusion 225
  echogram 221
  empirical models 215
  finite-difference time-domain 228
  geometrical models 220
  gypsum cavity wall absorption 229
  head-related transfer functions (HRTF) 232
  hybrid models 215, 230
  image models 220
  impulse response 232
  medium considerations 217
  modeling 215
  multichannel reproduction 233
  numerical model 227–228
  parametric room impulse response (PRIR) 234
  physical models 215
  propagation of sound waves 217
  ray-tracing models 223
  real-time auralization 233
  receiver 218
  reflectogram 221
  small room models 228
  sound transmission class (STC) 229
  source 218
  Statistical Energy Analysis 228
  steady-state waves 218
  surface materials 219
  thermal relaxation 217
  virtual reality 233
  wave equation 225
  wave equation models 225
Acoustical noise
  analog discs 1020
Acoustical noise control
  absorption coefficient 93
  acoustical doors 83
  air gap 79
  angle of incidence 72, 93
  attenuation of noise in ducts 91
  balanced noise criterion 68
  cavity resonance 86
  ceiling construction 78
  concrete block walls 77
  concrete walls 77
  continuous underlayment 80
  critical frequencies 72
  damping 77
  empirical mass law 71
  exceedance level 70
  floating floors 79
  floor and ceiling systems 79
  floor construction 78–79
  Gypsum wallboard 74
  heating, ventilating, and air con-

  ditioning (HVAC) 89
  high-loss frame walls 74
  homemade acoustical doors 83
  homemade acoustical windows 86
  identification of HVAC noise 90
  isolation mount systems 80
  isolation systems 74, 89
  location of HVAC equipment 90
  low-frequency response 69
  noise criteria 67
  percentile levels 70
  plenum noise attenuators 92
  proprietary acoustical doors 83
  proprietary acoustical windows 86
  proprietary silencers 93
  raised slab or jack-up system 81
  room-within-a-room 82
  single-number STC rating 73
  site noise survey 69
  site selection 69
  sound barrier 71
  sound level meter 67, 70
  sound lock corridor 83
  STC number 72
  TEF 20 67
  TEF 25 70
  transmission loss 71
  transmission loss of a compound barrier 87
  transmission loss of single glass plates 84
  transmission loss of spaced glass plates 85
  tuned stub noise attenuator 92
  vibration isolation 90
  wall caulking 77
  wall construction 74
  windows 84
Acoustical plasters
  acoustical treatment for indoor areas 107
Acoustical Society of America 12
Acoustical testing
  test and measurement 1607
Acoustical tiles
  acoustical treatment for indoor areas 107
Acoustical treatment for indoor areas
  absorption 97
  absorption coefficients of building materials and absorbers 115
  absorption ratings 101
  absorption testing 98
  acoustical absorption 97
  acoustical baffles 108
  acoustical banners 109
  acoustical diffusion 115
  acoustical plasters 107
  acoustical tiles 107
  applications of diffusion 120
  bass traps 110, 113
  carpet 107
  chairs absorption 108
  closed cell foams 105
  diaphragmatic absorbers 109
  diffraction 120
  diffuser testing 117
  diffusion 115, 117
  digital room correction (DRC) 121
  discrete absorbers 108
  draperies 108
  fire safety codes 121
  glass fiber panel 103
  gypsum wallboard (GWB) 111
  Helmholtz resonators 111
  impact of absorption 115
  impedance tube testing methods 100
  medium density fiber panels 103
  melamine foam panels 106
  membrane absorbers 109
  microperforated materials 112
  mineral fiber board 104
  natural cotton panels 105
  Noise Reduction Coefficient (NRC) 101
  nonreverberant spaces 114
  numerical diffusers 118
  open-cell foams 105
  people absorption 108
  perforated coverings 105
  perforated membrane absorber 111
  polyurethane foam panels 106
  porous absorbers 102
  Q factor 110
  quadratic residue diffuser (QRD) 118
  random diffusers 119
  reflection 120
  refraction 120
  resonant absorbers 109
  reverberation chamber 98
  reverberation time 114
  RT equations 114
  Sabine absorption coefficient 101
  scattering 117
  scattering coefficients 118
  Schroeder diffusers 115
  SONEX 106
  Sound Absorption Average (SAA) 101
Acoustics
  absorption of sound by air 205
  atmospheric absorption 204
  atmospheric effects outdoors 203
  impulse response 232
  outdoor sound
    distributed loudspeaker systems 210
    effects of air temperature 204
    effects of atmospheric absorption of acoustical energy 204
    effects of relative humidity 204
    effects of temperature inversion gradient 204
Acoustics for auditoriums and concert halls
  %Alcons 157, 159
  absorbed sound power 151
  Acoustic Control System (ACS) 193
  Active Field Control System (AFC) 194
  Alcons 156–157, 159
  ambiophony 191
  angular range of uniform irradiation 181
  architectural acoustics 190

articulation loss 156
assisted resonance 192
attenuation of maximum 181
audience absorption 186
auditoriums, congress centers 163
average sound energy density 151
axis-near reflection areas 179
balconies 173
bass ratio 151
broadband sound absorbers 189
CARMEN 196
ceilings 171
center time 154
circles 173
circular effect 179
clarity measure C80 159
classical concert halls 164
critical distance 151
curved boundary surfaces 170
definition measure C50 154, 158
degree of definition D 154
degree of diffusivity d 180
diffuse reflection 178
diffusivity ratio 181
directed (local) reflection 178
directional sound reflection 181
displacement of maximum 181
dissipation coefficient 183
echo criterion 154
eigenfrequencies 192
electronic architecture 190–191
elliptical effect 180
energy criteria 152
equivalent sound absorption area 165
ERES 191
flow resistance 184
galleries 173
general concert halls 164
geometrical reflection 178
Helmholz resonators 185, 192
hyperbolic effect 180
intelligibility of sentences 157
intelligibility of words 157
Interaural Cross-Correlation Coefficient (IACC) 153
LARES System 194
lateral efficiency LE 161
logatoms 157
mean reverberation time 149
modulation reduction factor m(F) 155
modulation transmission function (MTF) 155
multi-channel reverberation (MCR) 192
multi-genre theaters 163
multi-purpose halls 163
near field 151
opera houses 163
orchestra pit 177
panel resonances 184
parabolic effect 180
performance zone 168
planning fundamentals 162
platform consideration in concert halls 175
polygon ground plan 170
porous materials 184
RASTI 159
register balance measure BR 162
reverberance measure H 161
Reverberation on Demand System (RODS) 193
reverberation radius 151
reverberation time (RT60) 148, 165
reverberation time duration 148
room acoustics 190
room shape 167
room volume 165
room-acoustical criteria 147
room-acoustical planning 164
rooms with variable acoustics 164
scattering coefficient s 180
show theaters 164
site lines 173
sloping of tiers 173
smooth curved surfaces 179
sound absorbers 183
sound absorption coefficient 149, 183
sound coloration measures 160
sound delay systems 191
sound energy component 152
sound form 165
sound pressure level 153
sound reflection coefficient 183
sound-field-proportionate factors 152
spatial impression measure R 160
specific flow resistance 184
Speech Transmission Index (STI) 155
spoken-drama theaters 163
sports halls, gymnasiums 164
strength measure G 153
structure coefficient s 184
subjective intelligibility tests 156
System for Improved Acoustic Performance 194
task of variable acoustics 189
time criteria 148
transmission coefficient 183
trapezoidal ground plan 169
travel-time-based reverberation-time enhancing systems 191
variable acoustics 188
vario-acoustics 188

Active acoustics
  virtual systems 1455
Active balanced inputs
  sound system design 1305
Active balanced line driver
  integrated circuits 360
Active balanced line receiver
  integrated circuits 360
Active Field Control System (AFC)
  acoustics for auditoriums and concert halls 194
Active filters 711
  filters and equalizers 793
Active microphone preamplifier
  integrated circuits 350
Active release time constants
  consoles 903
Actuate time
  relays 374
AD524
  integrated circuits 350
AD636

  integrated circuits 347
Adaptive equalizers
  filters and equalizers 802
Adaptors
  wire and cable 417
ADAT 1443
  consoles 988
Advanced Encryption Standard (AES)
  Optical disc formats 1153
AES/EBU digital audio cable
  wire and cable 430
AES/EBU digital coaxial cable
  wire and cable 430
AES/EBU digital peak meter
  VI meters 1004
AES-10—MADI
  consoles 987
AES11 1444
AES-3
  consoles 986
AES3 1443, 1452
AES3 I/O 1447
AES3 receiver
  digital audio interfacing and networks 1466, 1472
AES3 transmitter
  digital audio interfacing and networks 1466
AES-42
  consoles 986
AES42-2006 1452
AGC
  optical disc formats 1143
AIFF files
  optical disc formats 1135
Aiming loudspeakers
  computer aided sound system design 1374
  sound system design 1263
Air flow management
  heatsinks 369
Air gap
  acoustical noise control 79
  audio transformers 279–280
Air gap length
  loudspeakers 608
Air mass
  loudspeakers 635
Albert L. Thuras 9
Alcons 1396, 1398, 1400, 1404, 1408
  acoustics for auditoriums and concert halls 156–157, 159
  computer aided sound system design 1339
  sound system design 1251, 1254, 1263, 1265–1267, 1272, 1274
Alessandro Volta 6
Alex Badmieff 12
Alexander Graham Bell
  fiber optics 451
Algorithm development
  DSP technology 1172
Algorithmic composition programs
  MIDI 1123
Aliasing
  delay 811
Alien crosstalk (ANEXT)
  wire and cable 416
Alignment
  computer aided sound system design 1378
Alkaline-manganese batteries 698
All pass filter
  filters and equalizers 796
Alloy-junction transistor
  solid state devices 325
All-pass amplifier
  delay 809
Alpha
  solid state devices 329
  transistors 329
ALS-1 interference predictions
  loudspeaker clusters 649, 651
Altec 11
Altec Lansing Corp 12
Aluminum extruded heat sinks 368
AM broadcast
  assistive listening systems 1551
Ambient microphones
  personal monitors 1431
Ambient temperature effect
  resistors 245
Ambiophony
  acoustics for auditoriums and concert halls 191
Ambrose J. Fleming 8
American Institute of Electrical Engineers 12
American National Standards Institute
  display technologies 1579
Americans with Disabilities Act (ADA)
  Assistive listening systems 1545
Ampere
  fundamentals of audio and acoustics 32
Ampere-turns (AT)
  relays 374
Amplification factor
  gain structure 1227
  tubes 312
Amplifier 703
  active band pass filter 714
  active filter 710–711
  active high pass filter 714
  active low pass filter 714
  automatic turn-on muting 720
  balanced input 711
  Bessel polynomial 713
  Bode diagram 705–706, 711
  Butterworth polynomial 712–713
  capacitor high-pass filter 720
  class A output 717, 721
  class AB + B output 721–722
  class AB output 717, 721
  class AB plus B output 717
  class B output 717
  class D output 717, 721, 726
  class G output 723
  class H output 723
  class TD output 726
  complementary symmetry monolithic Darlington power transistors 719
  constant voltage distribution 716
  current amplifier 703
  dc-to-dc converter 725
  EMI filter 724
  feedback theory 707
  field-effect transistors 718
  fuse 719
  group delay distortion 711

- high power 720
- inductor 713
- input-output impedance 703
- inverting operational amplifier 710
- inverting voltage amplifier 711
- Laplace transform 705
- linear 703
- load protection mechanism 719
- logarithmic 703
  - VU meters 1002
- magnetic field 703–704
- MOSFET transistors 720
- negative feedback 708–709
- noninverting voltage amplifier 711
- Nyquist diagram 705, 708
- operational amplifier 710–711, 713
- phase distortion 713
- pole-zero diagram 706
- power
  - fundamentals of audio and acoustics 25
- protection mechanisms 719
- switch mode power supplies 724–725
- thermal sensor 720
- totem pole with class AB plus B output 718
- transconductance 703
- transfer function 704
- transformer 720
- transresistance 703
- types of amplifiers 703
- voltage 703

Amplifier power ratings
- gain structure 1231

Amplifier sensitivity
- gain structure 1230

Amplifiers 1074
- differential 339
- inverting 338
- noninverting 338
- operational transconductance 340
- summing inverter 340

Amplitude
- units of measurement 1652

Amplitude response
- loudspeakers 628, 631

Amplitude spectra
- psychoacoustics 57

Amplitude spectrum
- psychoacoustics 50

Amplitude versus time
- test and measurement 1611

Analog A and C weighted filters
- what's the ear for? 1639

Analog delay 812

Analog discs
- acoustic feedback 1035
- brushes 1036
- cartridge voltage output 1033
- cleaning records 1037
- compliance 1032
- DIN curve 1017
- distortion 1017
- effective tonearm length 1023
- electrical loading 1033
- electrodynamic type cartridges 1027
- elliptical stylus 1031
- induced magnet cartridges 1028
- laser turntable 1034
- linear speed 1023
- magnetic field 1027–1028
- moving coil cartridges 1027
- moving magnet cartridges 1028
- NAB (RIAA) curve 1017
- phono preamplifiers 1033
- piezoelectricity 1030
- record care 1036
- record-cleaning machines 1036
- recording characteristics 1017
- semiconductor pickup cartridges 1029
- signal-to-noise ratio (SNR) 1016
- skating force 1023
- SNR 1033
- spherical stylus 1031
- step-up transformer 1027
- stereo disc groove 1031
- stroboscopic disc 1019–1020
- styli 1031
- stylus cantilever 1032
- stylus characteristics 1032
- stylus tip 1031
- tangential error 1022
- tangential tracking 1021
- tonearm antiskating force 1025
- tonearm geometry 1021
- tonearm resonant damping 1025
- tonearm vertical tracking angle 1025
- tonearms 1021
- turntable 1018
- turntable drive systems 1018
- turntable platter 1018
- turntable speed of rotation 1019
- variable reluctance cartridges 1027
- vertical resonance 1033

Analog display
- display technologies 1581

Analog record 1015

Analog shift register delays
- delay 810

Analog speakerphone
- preamplifiers and mixers 760

Analog to digital conversion
- delay 813

Analog to digital converter
- DSP technology 1167

Analog Z weighting (linear) filter
- what's the ear for? 1639

Analog-to-digital converter
- consoles 942

Analytical Model
- Full-Wave Methodology 226

Andre Marie Ampere 6

Anechoic chamber
- small room acoustics 141

Anechoic measurements
- test and measurement 1614

Angle of incidence
- acoustical modeling and auralization 222
- acoustical noise control 72, 93
- fiber optics 454–455, 479
- small room acoustics 139

Angle of radiation
- computer aided sound system design 1348

Angle of reflection
- fiber optics 455

Angle of refraction

  - fiber optics 454
- Angular directivity gain
  - computer aided sound system design 1356
- Angular directivity ratio
  - computer aided sound system design 1347, 1356
- Angular range of uniform irradiation
  - acoustics for auditoriums and concert halls 181
- Annoyance factor
  - what's the ear for? 1637
- ANSI S3.5 1997 1409
- Antenna cables
  - microphones 569
- Antenna splitters
  - microphones 570
- Antennas 560
  - assistive listening systems 1550
  - fundamentals of audio and acoustics 28
  - microphones 568
- Anti-alias filter
  - delay 812
- Anti-aliasing
  - delay 811
- Anti-aliasing filters
  - consoles 956
- Anti-imaging
  - delay 811
- Antiskating force
  - analog discs 1025
- Aperture time
  - delay 812
- Applications for personal monitors
  - Personal monitors 1425
- Applications of diffusion
  - acoustical treatment for indoor areas 120
- Architectural acoustics 190
  - fundamentals of audio and acoustics 28
- Arithmetic symetry
  - filters and equalizers 785
- Array height versus wavelength
  - loudspeaker clusters 659
- Array lobing
  - computer aided sound system design 1378
- Art Davis 14
- Arthur Edwin Kennelly 7
- Arthur Fleming 8
- Articulation Index (AI)
  - designing for speech intelligibility 1391, 1404, 1409
  - loudspeaker clusters 654
- Articulation loss
  - acoustics for auditoriums and concert halls 156
- Articulation loss of consonants
  - designing for speech intelligibility 1396
- Artificial ambience
  - sound system design 1328
- Artificial head systems
  - microphones 553
- ASHRAE 90
- Aspect ratio
  - display technologies 1581
- Assisted resonance
  - acoustics for auditoriums and concert halls 192
- Assistive Listening Devices (ALDs)
  - assistive listening systems 1545
- Assistive listening systems
  - AM broadcast 1551
  - Americans with Disabilities Act (ADA) 1545
  - antennas 1550
  - Assistive Listenig Systems (ALSs) 1545
  - Assistive Listening Devices (ALDs) 1545
  - carrier current antenna
    - Carrier current antenna assistive listening systems 1552
  - electronic system 1549
  - field distortion 1548
  - field strength 1547
  - FM broadcast 1550
  - hearing aid response 1546
  - hearing impaired 1545
  - induction loop system 1550
  - inductive reactance 1548
  - infrared control transmitter 1553
  - infrared link 1553
  - infrared slave emitters 1553
  - infrared systems 1552
  - infrared systems channel allocation 1555
  - infrared systems modulation technique 1555
  - IRLED transmitting diode 1554
  - large area multi-channel systems 1553
  - large area systems (LAS) 1553
  - light emitting diode (LED) 1553
  - light spectrum 1555
  - loop impedance 1548
  - lossy coaxial cable antenna 1552
  - Low Power Radio Services (LPRS) 1550
  - magnetic field strength 1546, 1548
  - magnetic induction loop system 1546
  - magnetic induction loops 1546
  - medium area systems (MAS) 1553
  - multiple-loop current flow 1548
  - National Association of the Deaf (NAD) 1545
  - normalized field strength 1547
  - open wire antenna 1552
  - personal listening systems 1553
  - sensitivity of the eye 1554
  - T-coil 1546
  - T-coil hearing aids 1551
  - Wheeler's equation 1549
  - wire data 1550
  - wireless IR system receiver 1557
- Assistive Listening Systems (ALSs)
  - Assistive listening systems 1545
- Astigmatism
  - optical disc formats

1138–1139
ASTM E336-97 72
Asymmetric-directivity horns
  loudspeakers 613
Asynchronous connections
  digital audio interfacing and networks 1461
Atmospheric absorption 204
  designing for speech intelligibility 1401
Attack time
  consoles 900
Attenuation
  fiber optics 458, 477, 479
  wire and cable 432
Attenuation coefficient
  fiber optics 458
Attenuation in fiber
  fiber optics 459
Attenuation measurement
  fiber optics 461
Attenuation of Maximum
  acoustics for auditoriums and concert halls 181
Attenuation of noise in ducts
  acoustical noise control 91
Attenuation of sound indoors
  sound system design 1248
Attenuation with increasing distance
  sound system design 1240
Attenuator
  fiber optics 479
Attenuator pads
  consoles 865
Attenuators
  attenuator networks 767
  automated faders 779
  automatic 779
  balanced attenuator 767
  balanced bridged T 774
  balanced ladder 776
  balanced T pad 773
  bridged T pad 773
  combining networks 772
  distortion 767
  dividing networks 772
  feedback-type volume control 778
  field-effect transistor attenuators 779
  H pad 773
  impedance 770
  impedance matching 769
  impedance mismatch ratio 770
  insertion loss 769, 773
  K factors for calculating attenuator loss 768
  K value 768
  L pads 771
  ladder 776
  ladder networks 776
  ladder potentiometers 776
  light-dependent attenuator (LDA) 778
  loss 769
  microphones 581
  minimum-loss pad 769
  mixer 780
  O type 775
  pi or delta attenuator 774
  summing amplifiers 781
  T attenuators 773
  transformer isolated 767
  types of attenuators 771
  U type 775
  unbalanced pad 767
  unity-gain current-summing amplifiers 781
  VI meters 1000
  voltage-controlled amplifier (VCA) 779
  volume and loudness controls 776
Atwater Kent 16
Audible frequency range
  units of measurement 1679
Audible spectrum
  fundamentals of audio and acoustics 27
Audience absorption
  acoustics for auditoriums and concert halls 186
Audio and video teleconferencing
  sound system design 1329
Audio archival
  message repeaters and evacuation systems 1527
Audio codecs
  optical disc formats 1152
Audio cooling tube
  heatsinks 369
Audio dB meter
  integrated circuits 348
Audio encoding
  psychoacoustics 48
Audio Engineering Society 13, 637
Audio frequency range
  units of measurement 1666
Audio over Ethernet
  consoles 989
Audio path
  consoles 855
Audio program material
  fundamentals of audio and acoustics 35
Audio signal transmission via computer networks
  sound system design 1333
Audio spectrum
  units of measurement 1678
Audio standards
  units of measurement 1663
Audio transformers
  air gap 279–280
  auto-transformer 280
  balanced input 297–298
  balanced input stages 302
  balanced output 298
  bandwidth 289
  basics 275
  B-H loop 278–279
  bi-directional reflection of impedances 287
  bi-filar winding 281
  bi-filar wound 286
  bobbin 282
  boosting 280
  bucking 280
  buzz 296
  cable capacitance 289
  capacitance 282
  center tapped 276
  CMRR 283, 290–291
  coil inductance 278
  common mode rejection ratio 290
  common-mode noise 283, 290
  common-mode rejection 302

condenser microphone 295
Co-Netic 283
constant-voltage 298
constant-voltage distribution 298
constant-voltage line 299
core 278
core laminations 279
core material 278, 284
coupling coefficient 276
damping 286
data sheets 303
demagnetizing 306
direct box 296
distortion 279, 283, 289
duplex 300
eddy currents 280
eddy-current losses 279
electronically-balanced inputs 302
electrostatic shield 282
excitation current 276–277, 280, 283
excitation flux 277
Faraday shield 282, 291–292
feedback factor 287
ferro-magnetic materials 278
flux 275
flux density 275, 278–279, 285
frequency response 285
gapping 279
generator impedance 285
ground noise 282
high-hysteresis materials 279
hum 296
HyMu 283
hysteresis 279, 284
hysteresis distortion 284, 306
hysteresis loop 278
IEC 60268-3 302
IM distortion 284–285
induction 275
inductive reactance 275–276, 285
inductor 275
input 289
input transformer 288
insertion loss 286, 290
installation 306
internal winding resistance 283
inter-stage coupling 293
isolation 290
law of induction 275
layering 281
leakage inductance 281
line input 291
line level shifter 303
line-level output 292
lines of force 275
load impedance 297
loudspeaker matching 300
magnet wire 282
magnetic circuits 280
magnetic coupling 288
magnetic field 275
magnetic flux 277, 294
magnetic flux density 284
magnetic hysteresis 279
magnetic operating point 279
magnetic saturation 283
magnetic shield 283, 301
magnetic shielding 283
magnetizing force 278
maintenance 306
measurements 302
microphone impedance conversion 295
microphone input 290
microphone output 294
microphonic 306
moving coil phono step-up 301
moving-coil phonograph pickup 291
multi-filar winding 281
Mumetal 283
negative feedback 287
negative output resistance 287
noise figure 289
non-magnetic materials 278
output 289, 296
output transformers 286, 290
parasitics 277
Permalloy 283
permeability 278
phase distortion 285
polarity 276
power transformer 280
repeat coil 300
residual field 279
resistance 280
ring core 280
saturate 278
series leakage inductance 286
signal transmission characteristics 302
single-ended power amplifier 294
SMPTE 284
source impedance 285, 287
step-down 276, 298
step-down auto-transformer 281
step-up 276, 298
telephone hybrid circuit 300
telephone isolation 300
testing 302
THD 284
toroidal core 280
toroidal power transformer 283
total harmonic distortion 284
transmission characteristics 302
turns ratio 276–277, 280, 299, 302
unbalanced input 297–298
unbalanced interface 296
unbalanced output 297–298
universal winding 282
unshielded output 283
voltage ratio 276
winding resistance 280
windings 276
zero rejection 296
zero-impedance amplifier 288
μ 278

Audio waveform fundamentals
  gain structure 1224
Audio with Video (AV) disc
  optical disc formats 1149
Audio/video receivers
  surround sound 1598
Audition
  consoles 827
Auditoriums, congress centers
  acoustics for auditoriums and concert halls 163
Auditory area
  psychoacoustics 50
Auditory filters
  psychoacoustics 49

Auditory nerve
  personal monitors 1433
Auditory system
  psychoacoustics 45
Audyssey Dynamic EQ
  surround sound 1600
AURA
  computer aided sound system design 1373
Auralization
  acoustical modeling and auralization 215, 231
  computer aided sound system design 1374
Auralization process
  acoustical modeling and auralization 232
Auricle
  psychoacoustics 43
Auto-focusing
  optical disc formats 1138
Automated faders 779
Automatic equalization
  sound system design 1312
Automatic gain control (AGC)
  preamplifiers and mixers 747
Automatic level control
  designing for speech intelligibility 1392
Automatic microphone mixers
  preamplifiers and mixers 741
Automatic microphone mixers implemented in software 752
Automatic microphone mixers with matrix mixing to multiple outputs 752
Automatic microphone mixing
  sound system design 1282
Automatic microphone muting
  consoles 905
Automatic microphone systems 540
Automatic mixing controller
  preamplifiers and mixers 752
Automatic mixing problems
  sound system design 1285
Auto-tracking
  optical disc formats 1139
Auto-transformer
  audio transformers 280
Auxiliary channel feeds
  consoles 930
Avalanche diode
  solid state devices 321
Avalanche photodiode (APD)
  fiber optics 479
Avalanche region
  solid state devices 320
Average absorption coefficient
  units of measurement 1660
Average level detection
  integrated circuits 345
Average power
  VI meters 1006
Average sound energy density
  acoustics for auditoriums and concert halls 151
Average voltage
  gain structure 1224
Average-responding level detection
  integrated circuits 345
Averaging
  consoles 962
Avery Fisher 5
A-weighting level
  psychoacoustics 52
Axial mode
  small room acoustics 128
Axial ray
  fiber optics 479
Axis-near reflection areas
  acoustics for auditoriums and concert halls 179

B

Back emf
  grounding and interfacing 1183
Backbone cabling
  wire and cable 412
Backscattering
  fiber optics 479
Backward masking
  psychoacoustics 48
Baffle
  loudspeakers 626
Balanced armature earphones
  personal monitors 1424
Balanced attenuators 767
Balanced bridge
  integrated circuits 353
Balanced bridged T attenuators 774
Balanced cable
  grounding and interfacing 1200
Balanced input
  amplifier 711
  audio transformers 297–298
  grounding and interfacing 1208
Balanced input circuits
  grounding and interfacing 1213
Balanced input stage
  audio transformers 302
  preamplifiers and mixers 740
Balanced interfaces
  grounding and interfacing 1212
Balanced ladder attenuators 776
Balanced line
  sound system design 1296
Balanced line input circuits
  integrated circuits 356
Balanced line interfaces
  integrated circuits 353, 355
Balanced line outputs
  integrated circuits 354
Balanced line receiver
  integrated circuits 356–357
Balanced lines
  wire and cable 407
Balanced microphones 521
Balanced mixing
  consoles 915
Balanced noise-criterion (NCB)
  acoustical noise control 68
Balanced output
  audio transformers 298
  grounding and interfacing 1208
Balanced output line amplifiers
  consoles 932
Balanced power
  grounding and interfacing 1215
Balanced T pad 773
Balconies
  acoustics for auditoriums and concert halls 173
Baluns
  wire and cable 416
Band limiting
  consoles 843
Band pass
  filters and equalizers 785, 791
Band pass filter 714

Band reject filter
  filters and equalizers 791
Bandpass filter
  consoles 884
Bandpass parametric equalizer
  consoles 969
Bandwidth 1388
  audio transformers 289
  consoles 864, 879
  fiber optics 468, 479
  filters and equalizers 785
  grounding and interfacing 1203
  preamplifiers and mixers 738
  relays 374
Bandwidth distance product
  fiber optics 479
Bar graph VU meter 1002
Bark scale
  psychoacoustics 49
Barrier block connectors
  sound system design 1293
Base current curves
  transistors 333
Base station microphone 530
Basic electronics
  grounding and interfacing 1181
Bass horns
  computer aided sound system design 1350
Bass ratio
  acoustics for auditoriums and concert halls 151
Bass reflex boxes
  loudspeakers 1350
Bass traps
  acoustical treatment for indoor areas 110, 113
Batteries
  absorbed glass mat (AGM) 696
  alkaline-manganese 698
  lead-acid 694
  lead-dioxide batteries 695
  LeClanche (Zinc-Carbon) 697
  mercury dry cells 699
  Nickel-Cadmium 697
  power supplies 692
Battery charging
  power supplies 693
Baxandall
  consoles 890
BD
  optical disc formats 1153
BD-ROM
  optical disc formats 1153
BD-Video
  optical disc formats 1153
Beam steering
  loudspeaker clusters 656, 662
Beam-power
  tubes 311
Beamsplitter
  fiber optics 479
Beam-tracing models
  acoustical modeling and auralization 224
Beamwidth
  fiber optics 479
Beat packets
  digital audio interfacing and networks 1505
Bel
  fundamentals of audio and acoustics 24
Bell Telephone Laboratories 11
  VU meters 998
Belt pack receiver
  personal monitors 1416
Beltpacks
  intercoms 1575
Bend radius fibers
  fiber optics 459
Beneficial destructive interference
  loudspeaker clusters 653
Beneficial interference
  loudspeaker clusters 652
BER (Bit Error Rate)
  fiber optics 480
Bessel filters
  consoles 890
Bessel function
  loudspeakers 634
Bessel polynomial 713
Beta
  solid state devices 329
  transistors 329
B-H loop
  audio transformers 278–279
Bias circuits
  transistors 329
Bias, magnetic
  relays 374
Biconic connector
  fiber optics 471
Bidirectional microphone 493, 495
Bi-directional reflection of impedances
  audio transformers 287
Bi-filar winding
  audio transformers 281
Biflex
  loudspeakers 598
Bifurcated contacts
  relays 377
Bimorph bender
  loudspeakers 604
Binaural
  psychoacoustics 57
Binaural hearing
  psychoacoustics 60
Binaural impulse response (BIR)
  acoustical modeling and auralization 232
Binaural pitches
  psychoacoustics 57
Binaural process
  psychoacoustics 56
Binaural recording microphones 553
Binaural reproduction
  acoustical modeling and auralization 232
  headphones 232
  loudspeakers 232
Biphase-mark
  digital audio interfacing and networks 1465
Bipolar field-effect transistors (BiFETs)
  consoles 837
Bipolar junction transistor
  integrated circuits 343
Biquad
  consoles 966
Biquad coefficient sets
  consoles 967
Biquad filter
  filters and equalizers 799
Biquad IIR filter
  filters and equalizers 799

Biradial horns
  loudspeakers 613
Birefringence
  display technologies 1587
Bistable magnetic latching relay 380
Bit stream corruption 1006
Bit-error-rate
  fiber optics 470
Blackman-Harris
  Test and measurement 1623
Blu-ray
  optical disc formats 1133
Blu-ray disc system
  optical disc formats 1153
Bobbin
  audio transformers 282
Bode diagram 705–706, 711
Bolt footprint
  small room acoustics 133
Bonded fin
  heatsinks 369
Bone conduction
  psychoacoustics 45
Bonello Criteria
  small room acoustics 132
Boosting
  audio transformers 280
Bootstrap
  consoles 881
Bounce
  relays 383–384
Bounce, contact
  relays 374
Boundary element analysis
  acoustical modeling and auralization 227
Boundary Element Method (BEM)
  Computer aided sound system design 1354
Bounded input/bounded output stable
  DSP technology 1164
Braid shields
  wire and cable 426
Breakdown region
  solid state devices 320
Breakdown voltage
  relays 374
  transistors 332
Bridged interfaces
  gain structure 1224
Bridged T pad 773
Brightness
  delay 808
Brilliant Color
  display technologies 1586
British Broadcast Standard
  VI meters 1004
Broadband sound absorbers
  acoustics for auditoriums and concert halls 189
Broadcast intercoms
  intercoms 1569
Broadcast matrix systems
  intercoms 1571
Broadcast party line system
  intercoms 1570
Bruel and Kjaer 9
Brushes
  analog discs 1036
Bucking
  audio transformers 280
Buckminster Fuller 5
Buffer
  fiber optics 480
Buffering latency
  consoles 990
Build-up of background noise and reverberation
  preamplifiers and mixers 743
Bundle identification
  digital audio interfacing and networks 1506
Burrus LED
  fiber optics 463, 480
Butterworth
  virtual systems 1449
Butterworth filter
  filters and equalizers 792
Butterworth filters
  consoles 890
Butterworth polynomial 713
Butterworth polynomials 712
Buzz
  audio transformers 296
B-weighting level
  psychoacoustics 52

C

C35 1388
  designing for speech intelligibility 1405
C50 1388
  designing for speech intelligibility 1404–1405
C7
  designing for speech intelligibility 1405
Cabinet diffraction
  loudspeakers 626
Cable
  1000Base-LH 1490
  1000Base-Lx 1490
  1000Base-SX 1490
  1000Base-T 1490
  100BaseT 1490
  10BaseF 1490
  10BaseT 1489, 1501
  Category 1 1498
  Category 2 1498
  Category 3 1498
  Category 4 1498
  Category 5 1498
  Category 5e 1498
  Category 6 1498
  Category 7 1498
  characteristics of typical fiber cables 457
  ethernet (Cat-5) 1430
  grounding and interfacing 1184
  MIDI 1111
  shield grounding
    grounding and interfacing 1213
  sound system design 1292
  specifications
    fiber optics 476
  types of fiber 456
Cable capacitance
  audio transformers 289
  grounding and interfacing 1187
Cable installation
  digital audio interfacing and networks 1503
Canadian Electrical Code cable substitution chart
  wire and cable 441
Candelas per meter squared
  display technologies 1579
Capacitance 253

  audio transformers 282
  grounding and interfacing 1182
  internal 313
Capacitance (C)
  units of measurement 1650
Capacitance of a shielded twisted pair cable
  preamplifiers and mixers 736
Capacitive coupling
  grounding and interfacing 1187
Capacitive loads
  relays 379
Capacitive networks
  filters and equalizers 787
Capacitive reactance 270
  consoles 876
  fundamentals of audio and acoustics 33
  grounding and interfacing 1182
Capacitor
  metal oxide-semiconductor 810
  metal-oxide-semiconductor junction FET capacitor 810
Capacitor filters
  power supplies 674
Capacitor microphone 507
Capacitor radio-frequency microphone 512
Capacitor radio-frequency, frequency-modulated microphone 511
Capacitors 251
  ac ripple current rating 258
  aluminum electrolytic 257
  capacitance 253
  capacitive reactance 270
  ceramic 257
  color codes 256
  dielectric absorption (DA) 255
  dielectric constant 252
  dissipation factor (DF) 255
  electrical characteristics 258
  electrochemical double-layer capacitors 262
  electrolytic 257
  equivalent series inductance (ESL) 255
  equivalent series resistance (ESR) 255
  film capacitor 256
  foil tantalum 260
  force of attraction 253
  impedance 258
  in ac circuits 251
  inductive reactance 258
  insulation resistance 255
  leakage current 258
  maximum working voltage 255
  mica 257
  network transfer function 253
  paper foil filled 257
  parallel plate capacitor equation 252
  power factor (PF) 255
  quality factor (Q) 255
  reverse voltage 259
  solid-electrolyte, sintered anode tantalum 261
  sound system design 1278
  standard life tests 259
  supercapacitors 262
  suppression 261
  surge voltage 259
  tantalum 259
  time constant 253
  types of capacitors 255
  ultracapacitors 262
  wet-electrolyte, sintered anode tantalum 260
  X type 261
  X Y type 262
  Y type 261
Capstan speed control 1045
Capture ratio
  intercoms 1575
Capture ratio and muting
  microphones 560
Carbon microphone 502
CARMEN
  acoustics for auditoriums and concert halls 196
Carpet
  acoustical treatment for indoor areas 107
Carrier-sense multiple access
  digital audio interfacing and networks 1488
Carry current
  relays 374
Cartridge voltage output
  analog discs 1033
Cartridges
  electrical loading 1033
  electrodynamic 1027
  induced magnet 1028
  magnetic field 1027–1028
  moving coil 1027
  moving-magnet 1028
  piezoelectricity 1030
  semiconductor pickup 1029
  step-up transformer 1027–1028
  styli 1031
  variable reluctance 1027
  voltage output 1033
Cascaded biquads
  consoles 968
Case temperature
  heatsinks 373
CAT5 connectors
  sound system design 1293
CAT-5/RJ-45
  consoles 989
Cat-5e cables
  interpretation and tour group systems 1531
Category 1 cable
  digital audio interfacing and network 1498
Category 2 cable
  digital audio interfacing and networks 1498
Category 3 cable
  digital audio interfacing and networks 1498
Category 4 cable
  digital audio interfacing and networks 1498
Category 5e cable
  digital audio interfacing and networks 1498
Category 6 cable
  digital audio interfacing and networks 1498
Category 7 cable
  digital audio interfacing and networks 1498
Category cable
  digital audio interfacing and networks 1498
Category cables

wire and cable 412
Cathode
tubes 311
CATV
grounding and interfacing 1208
CATV (Community Antenna TV) 439
CATV broadband cable
wire and cable 419
Causal system
DSP technology 1164
Cavity resonance
acoustical noise control 86
CCD
delay 810
Ccondenser microphone 582
CCTV (closed circuit television) cable
wire and cable 419
CD player hardware architecture
optical disc formats 1137
CD-DA
optical disc formats 1133
CD-R
optical disc formats 1133, 1142
CD-ROM
Optical disc formats 1133
optical disc formats 1141
CD-ROM/XA
optical disc formats 1142
CD-RW
optical disc formats 1133
CDRW
virtual systems 1439
CD-UDF
optical disc formats 1143
CD-WO (Write Once)
optical disc formats 1142
Ceiling construction
acoustical noise control 78
Ceiling loudspeaker systems
sound system design 1270
Ceilings
acoustics for auditoriums and concert halls 171
Cent
psychoacoustics 55
Center frequency
filters and equalizers 785
Center time
acoustics for auditoriums and concert halls 154
Central answering service
intercoms 1567
Central cluster loudspeaker
sound system design 1260
Central cluster plus distributed system
sound system design 1269
Central control unit
interpretation and tour group systems 1531
Ceramic microphone 503
Chairs absorption
acoustical treatment for indoor areas 108
Channel Coding
digital audio interfacing and networks 1465
Channel Definitions for Digital Cinema
surround sound 1601
Channel status data format
digital audio interfacing and networks 1467
Channel status format
digital audio interfacing and networks 1466
Channel-voice messages
MIDI 1104
Character
delay 808
Characteristic impedance
grounding and interfacing 1185–1186
VU meters 997
wire and cable 432–433, 437
Characteristic sound level
computer aided sound system design 1346
Charge coupled device
delay 810
Charge injection
consoles 940
Charge transfer device
delay 810
Charging voltage versus temperature
power supplies 694
Charles A. Coulomb 6
Charles Boner 14
Charles Proteus Steinmetz 6
Chebyshev filters
filters and equalizer 792
Chords
psychoacoustics 55
Chromatic dispersion
fiber optics 480
Chrominance
display technologies 1582
CIE Chromaticity Diagram (Color Chart)
display technologies 1580
Cilia
personal monitors 1433
CinemaScope
surround sound 1593
CIRC
optical disc formats 1135
CIRC code
optical disc formats 1147
CIRC decoder
optical disc formats 1140
CIRC encoding algorithm
Optical disc formats 1136
CIRC error correction
Optical disc formats 1141
Circles
acoustics for auditoriums and concert halls 173
Circuit theory
grounding and interfacing 1181
Circular effect
acoustics for auditoriums and concert halls 179
Cladding
fiber optics 480
Clarity measure C80
acoustics for auditoriums and concert halls 159
Class A circuitry
preamplifiers and mixers 740
Classical columns
computer aided sound system design 1350
Classical concert halls
acoustics for auditoriums and concert halls 164
Cleaning records

analog discs 1037
Clerk Maxwell 5
Clipping
gain structure 1225
Clocking
consoles 979
Closed box system
loudspeakers 636
Closed cell foams
acoustical treatment for indoor areas 105
CMOS
consoles 851
CMR
integrated circuits 354
CMRR 1208
Audio transformers 283
audio transformers 290–291
grounding and interfacing 1207, 1212
integrated circuits 340
preamplifiers and mixers 735, 740
Coarse wavelength division multiplexing
fiber optics 480
Coaxial cable
wire and cable 418
Coaxial cable construction
wire and cable 418
Coaxial cable installation
wire and cable 419
Coaxial cable signal loss
wire and cable 433
Coaxial cable termination
wire and cable 420
Coaxial loudspeaker 620
Coaxial shield
relays 374
CobraNet
digital audio interfacing and networks 1488, 1504
sound system design 1290
virtual systems 1443, 1447
CobraNet terminology
digital audio interfacing and networks 1504
Cochlea
personal monitors 1433
psychoacoustics 43
Cocktail party effect
psychoacoustics 60
Code development
DSP technology 1176
Coefficient compounding
consoles 979
Coefficient slewing
consoles 979
Coercivity
coercivityecording 1066
Coherent light
fiber optics 465, 480
Coherent-phase laser beam
optical disc formats 1138
Coil
relays 374
Coil AT
relays 374
Coil inductance
audio transformers 278
Coincident acoustical centers
loudspeaker clusters 650
Coincident microphone 542
Cold switching
relays 374
Collision detection
digital audio interfacing and networks 1488
Color code for nonpaired cables
wire and cable 407
Color code for paired cables
wire and cable 408
Color code for snake cables
wire and cable 409
Color code for UTP cable
wire and cable 415
Columbia Broadcasting System
VU meters 998
Column arrays
loudspeaker clusters 656
Column loudspeakers
loudspeaker clusters 661
Comb filter
delay 805
Comb filtering
personal monitors 1432
preamplifiers and mixers 743
Comb filters
designing for speech intelligibility 1390
microphones 521, 523
small room acoustics 139
Combination filters
power supplies 676
Combination shields
wire and cable 426
Combination tones
psychoacoustics 50
Combining networks
attenuators 772
Commercial intercom systems
intercoms 1566
Commercial security systems
intercoms 1564
Commissioning
sound system design 1292
Common impedance coupling
grounding and interfacing 1202
Common loudspeaker format
loudspeakers 629
Common mode rejection ratio
audio transformers 290
Common Mode Reverse Audio Feed (CMRAF)
interpretation and tour group systems 1534
Common-mode impedance
integrated circuits 354
Common-mode load impedance
integrated circuits 354
Common-mode noise
audio transformers 283, 290
Common-mode rejection
audio transformers 302
Common-mode rejection (CMR)
integrated circuits 354, 357
Common-mode rejection line receiver
integrated circuits 358
Common-mode rejection ratio (CMRR)
consoles 873
grounding and interfacing 1185, 1196
preamplifiers and mixers 739
Common-mode transfer
consoles 864
Common-mode voltage
grounding and interfacing 1185
Common-mode-rejection

consoles 863
Communication
loudspeakers 597
Communication protocols used in automatic microphone mixers
p 746
Community noise problems
what's the ear for? 1636
Compact Disc Digital Audio
optical disc formats 1133
Compact Disc Read Only Memory
optical disc formats 1141
Compact Disc Recordable (CD-R)
optical disc formats 1142
Companding
microphones 563, 567
Comparator elements
power supplies 681
Compensation
consoles 836
Complementary attenuators
consoles 829
Complex impedance
fundamentals of audio and acoustics 33
Complex power supplies 680
Complex tone
psychoacoustics 54
Compliance
analog discs 1032
stylus 1032
Component video
display technologies 1582
Composite video signal
display technologies 1582
Compound release time constants
consoles 902
Compressed function
psychoacoustics 52
Compression
consoles 906
Compression drivers
loudspeakers 600
Compression drivers and horns
sound system design 1256
Compression for effect
consoles 904
Compressors
personal monitors 1431
sound system design 1286
virtual systems 1439, 1445
Computational models
acoustical modeling and auralization 215, 220
Computational power
DSP technology 1176
Computer aided measurement systems
sound system design 1334
Computer aided sound system design
absorber data 1354
aiming loudspeakers 1374
Alcons 1339
alignment 1378
angle of radiation 1348
angular directivity gain 1356
angular directivity ratio 1347, 1356
array lobing patterns 1378
AURA 1373
auralization 1374
bass horns 1350
Boundary Element Method (BEM) 1354
characteristic sound level 1346
classical columns 1350
computer aided acoustic design (CAAD) 1339
computer model 1342
computer models for acoustic and sound system design 1339
cone tracing 1372
critical frequency 1363
data acquisition 1365
diffuse field 1346
diffuse-field sensitivity 1344, 1356
digitally controlled line arrays 1352
direct field 1346
directivity balloon curves 1343
directivity deviation ratio 1348
directivity factor 1346, 1349
directivity index 1349
DLL program module 1360
dummy head 1355
EASE SpeakerLab 1368
EASERA SysTune 1381
efficiency 1347
electro-acoustical transducers 1342
Finite Element Method (FEM) 1354
free-field open-circuit sensitivity 1356
free-field sensitivity 1344
frequency response. 1345
front to random factor 1346
front to random index 1346
generic loudspeaker library (GLL) 1362
head-related-transfer-functions (HRTF) 1355
horn loudspeakers 1350
human ear 1355
human voice 1343
image modeling 1371
line arrays 1352
loudspeaker arrays 1350
loudspeaker balloon measurements 1369
loudspeaker measurements 1357
loudspeaker polar data 1365
loudspeaker systems 1343
low-frequency absorption 1354
mapping 1378
measurement parameters for typical balloon resolutions 1358
medium-frequency horns 1350
microphones 1355
model-size dummy head 1341
natural sound sources 1343
nominal impedance 1345
phase interpolation 1364
point sources 1344
pressure open-circuit sensitivity 1356
pyramid tracing 1372
rated sensitivity 1346
ray radiation 1372
ray tracing 1367, 1371
reflectogram 1376
reverberation directional index 1349
room acoustic measures 1371
room acoustic simulation 1370
room impulse response 1340

  Room-Acoustic Analysis Module 1373
  scaled-model simulation 1341
  scale-model test 1340
  scattering data 1354
  sensitivity level 1345
  shadowing 1367
  simulation of complex loudspeakers 1360
  simulation of microphones 1369
  simulation of the human head 1369
  sound design 1374
  sound-field structure 1339
  SPL 1376
  time arrivals 1378
  tools of simulation 1370
  treble horns 1350
  wall materials 1354
  Wave Acoustics routines 1354
Computer aided system design
  sound system design 1333
Computer model
  computer aided sound system design 1342
Computer modelling
  designing for speech intelligibility 1400
Computer models for acoustic and sound system design
  computer aided sound system design 1339
Computer-Aided Acoustic Design (CAAD)
  computer aided sound system design 1339
Computers
  consoles 935
Computers and software for sound Reinforcement systems 1333
Concha
  psychoacoustics 43
Concrete block walls
  acoustical noise control 77
Concrete walls
  acoustical noise control 77
Condenser microphone 507, 533, 583
  audio transformers 295
condenser microphone 295
Conductance
  units of measurement 1650
Conduction
  heatsinks 367
Conduit fill
  wire and cable 431
Cone loudspeaker enclosures
  sound system design 1255
Cone loudspeakers
  sound system design 1255
Cone of confusion
  psychoacoustics 58–59
Cone radiator
  loudspeakers 610
Cone tracing
  computer aided sound system design 1372
Co-Netic®
  audio transformers 283
Cone-type loudspeakers
  loudspeakers 603
Conference bridge
  preamplifiers and mixers 758
Conference line systems
  intercoms 1561
Conference room ceiling systems
  preamplifiers and mixers 761
Conference room noise
  preamplifiers and mixers 755
Conference room reverberation
  preamplifiers and mixers 756
Conference room sound system
  sound system design 1328
Configuring a personal monitor system 1420
Connectors
  Biconic fiber optic 471
  couplers fiber optic 475
  D4 fiber optic 472
  Digital audio interfacing and networks 1499
  FC/PC fiber optic 471
  FDDI fiber optic 472
  fiber optic 1502
  installation
    fiber optics 473
  LC fiber optic 472
  MT-RJ fiber optic 473
  OpticalCon fiber optic 473
  SC fiber optic 472
  SFF fiber optic 472
  SMA 471
  Sound system design 1293
  ST fiber optic 472, 483
Console control surfaces
  consoles 949
Consoles
  1/F noise 861
  16 track routing 854
  32 track routing 854
  A/D conversion 942, 954
  absolute phase 891
  active release time constants 903
  ADAT 988
  AES-10–MADI 987
  AES3 986
  AES-42 986
  analog-to-digital (A/D) converter 942
  anti-aliasing filters 956
  attack time 900
  attenuator pads 865
  audio over Ethernet 989
  audio path 855
  audition 827
  automatic microphone muting 905
  auxiliary channel feeds 930
  averaging 962
  balanced mixing 915
  balanced output line amplifiers 932
  band limiting 843
  bandpass filter 884
  bandpass parametric equalizer 969
  bandwidth 864, 879
  Baxandall 890
  Bessel filters 890
  bipolar field-effect transistors 837
  biquad 966
  biquad coefficient sets 967
  bootstrap 881
  buffering latency 990
  Butterworth filters 890
  capacitative reactance 876
  cascaded biquads 968

CAT-5/RJ-45 989
changing filter frequency 878
charge injection 940
clocking 979
CMOS 851
coefficient compounding 979
coefficient slewing 979
common-mode rejection 863
common-mode rejection ratio (CMRR) 873
common-mode transfer 864
compensating operational amplifiers 841
compensation 836
complementary attenuators 829
compound release time constants 902
compression 906
compression for effect 904
computers 935
constant bandwidth 883
constant-amplitude phase-shift (CAPS) 883
constant-amplitude phase-shift network 883
control surface intelligence 951
control surfaces 948–949
control voltage noise 938
control-room monitor loudspeaker 826
convertor limitations 955
convertor requirements 955
convertors 954
crosstalk 853, 919
current limiting 887
damping factor 888
data acquisition 942
DAW internal audio paths 984
DAW limitations 985
dc-blocking series capacitors 868
debouncing 926
delay 961
delay line 966
design developments 833
desktop loudspeakers 918
devolved mixing 914
digital 953
digital audio networking 989
digital audio systems 954
Digital Audio Workstations (DAWs) 953, 981
digital control of outboard gear 953
digital mixer 976
digital mixer architecture 975
digital signal processors (DSP) 959
digitally controlled amplifiers 940
digital-to-analog (D/A) conversion 957
DirectX 985
discrete logic 942
discrete operational amplifiers 838
distortion 923
DSP filtering 963
dynamics processing 895
early drum recording 826
earth ground 846
echo 828, 962
effects send 828
electronic balanced outputs 869
electronic switching 851
equalization 875
equalization control 890
equalizer 875, 894
equalizer high-frequency response anomalies 970
equalizer response curves 876
expandor 906
expansion 907
fader automation 936
faders 946
fake inductor 883
Faraday shield 863
feed forward VCA-style dynamics 907
feedback phase-leading stability compensation 839
feedback-stye compressors 973
feedback-stye limiting 973
feed-forward sidechain 905
Field Programmable Gate Array (FPGA) 976
field-effect transistors 851
filter gain 888
filter quantization distortion 967
FIR coefficient sets 967
FIR filters 963
flash conversion 956
floating point processors 960
foldback 828
front end instability 842
function modes 923
gate attack 905
gating 905, 973
Gilbert Cell 937
ground current summing 848
ground noise 849
grounding 844, 913
grouping 831
gyrator 880
gyrator types 880
hang time 903
headroom 933
high-frequency control 894
high-pass filters 875, 888
hold time 903
host micro controller 981
IIR filters 966
impulse response 963
inadequate samples for limiters 974
inductance 879
inductive reactance 841, 876, 879, 888
input coupling 872
input impedance 865
input saturation 842
inputs 976
instability 838
instruction cycles 959
internal grounding 847
interpolation filter 958
inverse frequency noise 860
inverting amplifiers 869
jackfields 952
keying 906
ladder DACs 957
latency 991
level specifications 834
limiter 874
limiting 895
limiting for effect 904
line level inputs 867, 873
log-antilog cell 938
logic control 923

logic sense 928
loop filters 886
loop networking 992
low-frequency control 894
master functions 932
matrix crosstalk 854
mechanical recording 826
metering 980
meters 919, 933
microphone transformers 861
MIDI sequencing 983
minimizing noise 853
minimum gain 874
mix code 977
mix stage 976
mixdown mode 923
mixing 910
monitoring 826, 917
MOSFETs 851
motorized pots 946
moving digital audio around 985
multiband processing 902
multiple effect feeds 830
multiple memory spaces 961
multiplexers 945
multiplier/accumulator (MAC) 959
multitrack 922
multitrack routing 930
near-field loudspeakers 918
noise characteristics 887
noise figure 860
noise sources 910
nonmaskable interrupt 929
null indicators 946
open-circuit stability 864
operational amplifiers 833, 835
optimum source impedance (OSI) 861
output impedance 840
overdub mode 923
overdubbing 828
oversampling 958
panning 829
panpots 915
parallelism 961
parametric equalizer 968
passive mixing 914
peak program meter (PPM) 920
phase margin 838
phase shift 878
photoresistor 874
plug-ins 984
positive feedback 838
potentiometric switching 853
power transfer 859
pre-delays 975
prefade listen (PFL) 827, 918
processor control 855
processor types 960
programmable gate arrays 942
PROM 928
propagation delay 912
Q 879, 883, 888
radio-frequency inductors 912
reactance 878
reactive ground effects 850
recall display 945
reconstruction filters 958
recording mode 923
recursive processing 965
related crosstalk 919
relays 850
release time 901
resetting functions 946
resolution 954
resolvers 947
resonance 879
resonant Q 879
reverberation 828, 962
rms detection 908
router mixing 993
routing 850, 952
Sallen-Key filter 888
sample-rate convertors 958
Sample-Rate Convertors (SRCs) 976
sampling rate 955
second-order filters 890
serial audio format 976
serial networking 992
sidechain sources 909
sidechain time constants 900
sigma-delta conversion 957
sigma-delta convertors 956
sigma-delta DACs 958
signal switching 850
simulated resonance 883
single order networks 876
single-channel console 948
slew-rate effects 844
slew-rate limitations 836
solo, solo-free 918
stability 887
star network topology 992
stereo 829
streaming compression codecs 975
subgrouping 833
subtractive feedback gate 906
successive-approximation encoder 956
surround panning 916
swinging input control 892
swinging input equalizer 893
swinging output control 892
switch debouncing 926
symmetrical FIRs 965
system level architecture 871, 933
talkback 829
termination impedance 859
thresholding 908
time manipulation 961
time processing in DSP 961
time-domain effects 839
totem pole 880
transformer characteristics 862
transformerless front ends 866
transformers 867
transient intermodulation distortion 840
transistors 834
transversal filters 963
two-integrated loop 885
universal mix buses 978
unrelated crosstalk 919
USB 988
USB microphones 989
User Defined Protocol (UDP) 991
VCA channel application 939
Virtual Studio Technology 984
virtual systems 1440, 1445, 1452
virtual-earth mixers 910
virtual-earth mixing buses 932
voltage controlled amplifier

    (VCA) 936
  voltage controlled-attenuator 900
  voltage-controlled amplifier (VCA) 936
  VU meter 920
  wake-up time 905
  Wein Bridge 885
  winding resistance 883
  windowing 964
  wire 845
  zero impedance 834
  zero-crossing 941
Constant bandwidth
  consoles 883
Constant current interface
  gain structure 1224
Constant current regulated power supply 679
Constant directivity horns
  loudspeakers 612–613
Constant linear velocity (CLV)
  optical disc formats 1134
Constant Q equalizers
  sound system design 1287
Constant Q filters
  filters and equalizers 802
Constant voltage
  audio transformers 298
  fundamentals of audio and acoustics 25
  gain structure 1223
Constant voltage crossover
  loudspeakers 624
Constant voltage distribution
  audio transformers 298
Constant voltage interface
  gain structure 1223
Constant-amplitude phase-shift (CAPS) 883
Constant-amplitude phase-shift network
  consoles 883
Constant-k filters
  filters and equalizers 789
Constant-voltage line
  audio transformers 299
Constant-voltage regulated power supplies 678
Constant-voltage, constant-current regulated power supplies power supply 679
Contact
  relays 374
Contact arrangements
  relays 378
Contact bounce
  relays 378
Contact characteristics
  relays 377
Contact current
  relays 377
Contact resistance
  relays 383–384
  resistors 250
Contact Resistance Stability (CRS)
  relays 381
Contact resistance, dynamic
  relays 374
Contact resistance, static
  relays 374
Content Protection for Pre-Recorded Media (CPPM)
  optical disc formats 1150
Continuous output power
  gain structure 1231
Continuous to discrete conversion
  DSP technology 1168
Continuous underlayment
  acoustical noise control 80
Continuous-time signal
  DSP technology 1161, 1167, 1169
Contralateral ear
  psychoacoustics 58
Contrast
  display technologies 1579
Control grid
  tubes 311, 314
Control room
  small room acoustics 140
Control surface intelligence
  consoles 951
Control surfaces
  consoles 948
Control voltage noise
  consoles 938
Controlled-reluctance microphone 530
Controls and features of automatic microphone mixers
  preamplifiers and mixers 745
Convection
  heatsinks 367, 370
Conventional array shortcomings
  loudspeaker clusters 647
Conversion factors 1666
Converters
  power supplies 690
Convertor limitations
  consoles 955
Convertor requirements
  consoles 955
Convertors
  consoles 954
Convex surface
  fundamentals of audio and acoustics 29
Convolution
  DSP technology 1165
Cooper Sound CS104 four-channel ENG mixer
  preamplifiers and mixers 740
Core
  audio transformers 278
  fiber optics 480
Core laminations
  audio transformers 279
Core material
  audio transformers 278, 284
Core switching
  digital audio interfacing and networks 1498
Corner frequency
  filters and equalizers 785
Cosine sequence
  DSP technology 1163
Cosine signal
  DSP technology 1162
Cosine wave
  fundamentals of audio and acoustics 30
Coupler
  excess loss in fiber optics 475
  fiber optics 480
Coupling coefficient
  audio transformers 276
Coupling loss
  fiber optics 460, 480
Coverage 1399, 1454

Coverage angle
  loudspeaker clusters 648
Coverage boundaries
  loudspeaker clusters 649
CPPM
  optical disc formats 1150
Crest factor
  fundamentals of audio and acoustics 36
  gain structure 1224–1225, 1231
  VU meters 999
Critical angle
  fiber optics 480
Critical bands
  psychoacoustics 49
Critical bandwidths of the human ear
  psychoacoustics 49
Critical distance
  acoustics for auditoriums and concert halls 151
  preamplifiers and mixers 757
  psychoacoustics 61
  sound system design 1247
  test and measurement 1619
  units of measurement 1658
Critical distance modifier (M)
  units of measurement 1659
Critical frequencies
  acoustical noise control 72
Critical frequency
  computer aided sound system design 1363
Cross Interleave Reed Solomon Code
  optical disc formats 1135
Crossover
  virtual systems 1449
Crossover frequency
  loudspeakers 623
Crossover network
  loudspeakers 622
Crossover networks
  sound system design 1275
Crossover systems
  delay 809
Crosspoint electronics
  intercoms 1572
Crosstalk
  consoles 853, 919
  wire and cable 437
Crosstalk (crosstalk coupling)
  relays 375
Crystal microphone 503
CSL LightSplice System
  fiber optics 474
CTD
  delay 810
Cueing messages
  MIDI 1129
Current (I)
  units of measurement 1648
Current amplifier 703
Current flow
  transistors 328
Current limiting
  consoles 887
  power supplies 681
Current ratings
  wire and cable 403
Current-regulator circuit
  power supplies 680
Curved boundary surfaces
  acoustics for auditoriums and concert halls 170
Custom molded earphones
  personal monitors 1423
Custom sleeve earphones
  personal monitors 1423
Cut-back method test
  fiber optics 461
Cutoff frequency
  filters and equalizers 785
  loudspeakers 638
Cutoff wavelength
  fiber optics 480
C-weighting level
  psychoacoustics 52
Cyclic redundancy check (CRC)
  digital audio interfacing and networks 1489

D

D'Arsonval movement 998
D/A converters
  optical disc formats 1137
D/R ratio
  Designing for speech intelligibility 1388
  designing for speech intelligibility 1395
D4 fiber optic connector 472
Daisy chain
  MIDI 1112
Damping
  acoustical noise control 77
  audio transformers 286
  loudspeakers 599
Damping factor
  consoles 888
  filters and equalizers 793
  small room acoustics 131
  wire and cable 437
Damping materials
  acoustical modeling and auralization 228
Dante
  digital audio interfacing and networks 1511
Dark current
  fiber optics 480
Darlington circuit
  relays 390
DARS
  digital audio interfacing and networks 1463
DASH recorders
  magnetic recording 1079
DAT
  optical disc formats 1135
Data acquisition
  computer aided sound system design 1365
  consoles 942
Data buffers
  DSP technology 1174
Data cable
  10Base2 1489
  10Base5 1489
  10Broad36 1489
  1Base2 1489
Data decoding
  optical disc formats 1140
Data encoding
  optical disc formats 1134
Data rate
  fiber optics 480
Data sheets
  audio transformers 303
Data windows

test and measurement 1622
DAW internal audio paths
consoles 984
DAW limitations
consoles 985
dB
Audio and Acoustic DNA 10
dBA
VI meters 997
dBC
VI meters 997
what's the ear for? 1639
dBf
VI meters 997
dBm
VI meters 997
dBr
VI meters 997
dB-SPL
VI meters 997
dBu
sound system design 1299
VI meters 997
dBV
sound system design 1300
VI meters 997
dBW
VI meters 997
dc blocking series capacitors
consoles 868
Dc loads
relays 379
dc power supplies 672
DDL program module
computer aided sound system design 1360
DDP
optical disc formats 1134
Debouncing
consoles 926
Decay rate
delay 808
Decibel (dB)
fiber optics 480
units of measurement 1655
Decibel conversion
fundamentals of audio and acoustics 24
Decibel formulas
fundamentals of audio and acoustics 26
Decibel references
fundamentals of audio and acoustics 25
Decimation
delay 814
Decimation filter
optical disc formats 1144
Decoder
psychoacoustics 48
Dedicated line
intercoms 1561
Definition measure C50
acoustics for auditoriums and concert halls 154, 158
Definitions
wire and cable 412
Degree of definition
acoustics for auditoriums and concert halls 154
Degree of diffusivity d
acoustics for auditoriums and concert halls 180
Delay 805
acoustic 809
aliasing 811
all-pass amplifier 809
analog 812
analog shift register 810
analog to digital 813
anti-alias filter 812
anti-aliasing 811
anti-imaging 811
aperture time 812
brightness 808
character 808
charge coupled device 810
charge transfer device 810
comb filter 805
consoles 961
CTD 810
decay rate 808
decimation 814
Delta Modulation 813
delta modulator 813
digital 810
digital shift registers 811
distortion 813
DSP 811
dynamic range 810
Energy Time Curve 808
envelopment 808
flange 809
flutter echo 808
idling noise 814
Initial Time Gap 808
metal oxide-semiconductor capacitor 810
metal-oxide-semiconductor junction FET capacitor 810
MOSC 810
MOSJC 810
most significant bit 813
Nyquist frequency 811, 814
over-sampling 812, 814
perceived room size 808
phase shift 809
Pulse Code Modulation 813
quantization noise 813–814
random access memory 811
reticulating path 809
reverberation 808
reverberation synthesis 808
sample and hold 812
sampling 811, 813
Sigma-Delta Modulation 814
Sigma-Delta Modulator 814
SNR 810, 814
sound system design 1288
tape 810
tapped delay 809
virtual systems 1451
wire and cable 431
Delay line
consoles 966
Delay skew
wire and cable 416, 431
Delayed impulses
DSP technology 1164
Delta Modulation
delay 813
Delta modulator
delay 813
Demagnetizing
audio transformers 306
demagnetizingrding 1064
Dennis Gabor 18
Dense wavelength division multiplexing
fiber optics 480

Derived quantities in physics 1648
Design objectives for automatic microphone mixers
  preamplifiers and mixers 744
Design procedures
  fiber optics 477
Designation and Environmental Areas 439
Designing a complex cluster
  sound system design 1265
Designing a distributed ceiling system
  sound system design 1271
Designing for speech intelligibility
  %Alcons 1391, 1396, 1398, 1400
  absorption coefficient 1396
  Articulation Index (AI) 1391, 1404, 1409
  articulation loss of consonants 1396, 1404
  atmospheric absorption 1401
  automatic level control 1392
  C35 1405
  C50 1404–1405
  C7 1405
  comb filters 1390
  computer modeling 1400
  D/R ratio 1388, 1395
  direct + early reflections 1394
  direct sound 1394
  direct to reverberant measurements 1404
  direct to reverberant ratio 1393, 1396
  discrete echoes 1394
  distributed ceiling loudspeaker system 1395
  distributed loudspeaker systems 1398
  early reflections 1388, 1394
  early to late ratios 1404
  echo disturbance 1398
  echoes 1397
  equalization 1389, 1400
  intelligibility 1397
  intelligibility criteria 1403
  intelligibility measurement 1403
  intelligibility prediction 1395, 1400
  late reflections 1394
  loudness 1390
  loudspeaker
    directivity (Q) 1396
  loudspeaker directivity 1396
  Modified Rhyme Test (MRT) 1403
  nature of speech 1387
  noise sensing 1392
  Peutz equations 1396
  phonemically balanced (PB) word lists 1403
  polar ETC measurements 1409
  RASTI 1405
  RaSTI 1406–1407
  rate of delivery 1401
  reverberation 1394
  reverberation time (RT60) 1393, 1396
  signal to noise ratio 1390
  SII Speech Intelligibility Index 1409
  sound field components 1394
  spectator sports 1392
  speech intelligibility 1387
  Speech Interference Level 1403
  speech spectrum 1388
  Speech Transmission Index (STI) 1392, 1405
  Speech Transmission Index—STI 1397
  speech waveforms 1387
  STI 1405
  STI modulation reduction factor 1408
  STI/RaSTI 1408
  STIPA 1405, 1407
  STIPa 1407
  system distortion 1388
  talker articulation 1401
  U50 1405
  U80 1405
designing for speech intelligibility
  intelligibility optimization 1402
Desktop loudspeakers
  consoles 918
Destructive acoustic interference
  loudspeaker clusters 652
Detector
  fiber optics 480
Detector specifications
  fiber optics 476
Detectors
  fiber optics 467
Devolved mixing
  consoles 914
Diacs
  solid state devices 325
Diaphragm types
  loudspeakers 598
Diaphragmatic absorbers
  acoustical treatment for indoor areas 109
Dielectric absorption (DA) 255
Dielectric constant
  capacitors 252
Dielectric constants
  wire and cable 403, 425
Differential amplifier
  grounding and interfacing 1185
Differential amplifiers
  integrated circuits 339
Differential noise-canceling microphone 530
Differential polar relay 381
Differential signal
  integrated circuits 353
Diffraction
  acoustical treatment for indoor areas 120
Diffraction of the head
  psychoacoustics 58
Diffuse field
  computer aided sound system design 1346
Diffuse reflection
  acoustics for auditoriums and concert halls 178
Diffuse-field equalization
  microphone 555
Diffuse-field sensitivity
  computer aided sound system design 1344, 1356
Diffuser testing
  acoustical treatment for indoor areas 117
Diffusion

  acoustical modeling and auralization 225
  acoustical treatment for indoor areas 115, 117
Diffusivity ratio
  acoustics for auditoriums and concert halls 181
Digital audio
  digital audio interfacing and networks 1461
  moving it around 985
Digital audio connectors
  sound system design 1296
Digital audio devices
  personal monitors 1431
Digital audio interconnections
  digital audio interfacing and networks 1462
Digital audio interfacing and networks
  1000Base-LH 1490
  1000Base-LX 1490
  1000Base-SX 1490
  1000Base-T 1490
  100BaseT 1490
  100Base-TX 1490
  10Base2 1489
  10Base5 1489
  10baseF 1490
  10BaseT 1489
  10Broad36 1489
  1Base2 1489
  AES3 receiver 1466, 1472
  AES3 transmitter 1466
  asynchronous connections 1461
  beat packets 1505
  biphase-mark 1465
  bundle identification 1506
  cable installation 1503
  carrier-sense multiple access 1488
  Channel Coding 1465
  channel status data format 1467
  channel status format 1466
  CobraNet 1488, 1504
  CobraNet terminology 1504
  collision detection 1488
  connectors 1499
  core switching 1498
  cyclic redundancy check (CRC) 1489
  Dante 1511
  DARS 1463
  digital audio 1461
  digital audio interconnections 1462
  digital audio reference signal (DARS) 1463
  Digital Phantom Power 1467, 1474
  digital to analog converter 1474
  EMI 1503
  Ethernet 1487
  Ethernet equipment 1490
  Ethernet packet format 1488
  EtherSound 1510
  Fast Ethernet 1490
  fast link pulses (FLPs) 1494
  fiber optic cable 1502
  fiber optic cable bend radius 1503
  fiber optic connectors 1502
  frame check sequence (FCS) 1489
  Gigabit Ethernet 1490
  intrinsic jitter 1471
  isochronous connections 1462
  isochronous data packet 1505
  jitter 1471
  least significant bit (LSB) 1464
  line driver 1471, 1473, 1486
  line receivers 1471, 1473
  look up table (LUT) 1493
  managed switches 1495
  media access control (MAC) 1488
  media converter 1495
  meshing 1497
  microphone identification 1478
  most significant bit (MSB) 1464
  multicast bundles 1506
  multichannel audio digital interface 1484
  multimode fiber optic cable 1490, 1502
  normal link pulses (NLPs) 1494
  NRZI encoding 1485
  optical cable pull force 1503
  optical interface 1483
  PCM 1463
  phase comparator 1478
  phase locked loop 1474
  Power over Ethernet 1490
  preambles 1465
  private bundles 1507
  protocol 1505
  pulse code modulation 1463
  Quality of Service (QoS) 1496
  RealAudio 1461
  remote control pulses 1477
  reservation packets 1505
  RFI 1503
  router 1496
  routing 1496
  sample frequencies 1485
  Sample Rate Converter (SRC) 1463
  shielded twisted pair 1470
  Simple Network Management Protocol (SNMP) 1495
  single-mode fiber optic cable 1490, 1502
  Soundweb 1486
  spanning tree 1496
  SRC 1463
  Star of Stars topology 1491
  status flags 1478
  subframe format 1464
  switches 1493
  synchronization 1474, 1477, 1486
  synchronous connections 1461
  txUnicastMode 1506
  unbalanced electrical interface 1483
  unicast bundles 1506
  unshielded twisted pair 1470
  Virtual Local Area Network (VLAN) 1496
  voltage controlled crystal oscillator 1474
  Windows Media Audio (WMA 1461
  XLD connector 1481
Digital audio networking
  consoles 989
  sound system design 1290

Digital audio processors
  virtual systems 1445
Digital audio reference signal (DARS)
  digital audio interfacing and networks 1463
Digital audio storage
  message repeaters and evacuation systems 1515
Digital audio systems
  consoles 954
Digital Audio Workstations (DAWs)
  consoles 953, 981
Digital circuit terminology 361
Digital consoles
  consoles 953
Digital Content Protection
  display technologies 1582
Digital delays
  delay 810
Digital display
  display technologies 1581
Digital dynamics
  consoles 970
Digital echo canceller
  sound system design 1329
Digital filters
  filters and equalizers 797
Digital fixed matrix display scan conversion
  display technologies 1582
Digital integrated circuits 360
Digital Light Processing (DLP)
  display technologies 1585
Digital limiter
  consoles 972
Digital matrix systems 1562
  intercoms 1561, 1572
Digital Micromirror Device (DMD)
  display technologies 1585
Digital mixer
  consoles 976
Digital mixer architecture
  consoles 975
Digital on-hold players
  message repeaters and evacuation systems 1519
Digital peak program meter
  VI meters 1004
Digital Phantom Power
  digital audio interfacing and networks 1467, 1474
Digital potentiometers
  intercoms 1573
Digital recording
  magnetic recording 1083
Digital room correction (DRC)
  acoustical treatment for indoor areas 121
Digital sampling
  test and measurement 1620
Digital shift registers
  delay 811
Digital signal processing
  DSP technology 1161
  optical disc formats 1140
  sound system design 1288
Digital Signal Processor
  DSP technology 1172
Digital signal processor
  magnetic recording 1086
Digital signal processors
  message repeaters and evacuation systems 1515
Digital signal processors (DSP)
  consoles 959
Digital tape head
  magnetic recording 1083
Digital Theater Sound (DTS)
  surroound sound 1595
Digital to analog converter
  digital audio interfacing and networks 1474
Digital video
  wire and cable 422
Digital Video Interface (DVI)
  display technologies 1582
Digital Visual Interface (DVI)
  wire and cable 409
Digital wireless transmission
  microphones 572
Digitally controlled amplifiers
  consoles 940
Digitally controlled analog devices
  virtual systems 1440
Digitally controlled digital audio devices
  virtual systems 1441
Digitally controlled line arrays
  computer aided sound system design 1352
Digitally steered column arrays
  loudspeaker clusters 656
Digitally-steerable loudspeaker column arrays
  loudspeaker clusters 653
Digital-to-analog (D/A) conversion
  consoles 957
DIN curve
  analog discs 1017
Diode
  avalanche 321
  avalanche photodiode (APD)
    fiber optics 467, 479
  diacs 325
  laser
    fiber optics 481
  light emitting
    fiber optics 481
  noise 322
  photodiode
    fiber optics 482
  Shockley 322
  small-signal 321
  solid state devices 320
  tubes 311
  tunnel 322
  varactor 322
  zener 321
Diode versus synchronous rectifiers
  power supplies 685
DIP relays 374
Dirac delta function
  loudspeakers 640
Direct + Early Reflections
  Designing for speech intelligibility 1394
Direct box
  audio transformers 296
Direct field
  computer aided sound system design 1346
Direct radiator
  loudspeakers 610
Direct radiator enclosure design
  loudspeakers 635
Direct radiator loudspeaker 626
Direct sound
  designing for speech intelligibil-

ity 1394
  sound system design 1264
Direct sound radiation
  loudspeakers 633
Direct Stream Digital
  optical disc formats 1133
Direct Stream Digital (DSD)
  optical disc formats 1144
Direct Stream Digital (DSD) coding
  optical disc formats 1144
Direct to reverberant measurements
  designing for speech intelligibility 1404
Direct to reverberant ratio
  designing for speech intelligibility 1393, 1396
Direct versus reverberant field
  microphones 520
Direct/reverberant ratio
  sound system design 1246
Direct-coupled solid-state relays 390
Directed (local) reflection
  acoustics for auditoriums and concert halls 178
Direction dependent automatic microphone mixers 750
Directional energy radiation
  fundamentals of audio and acoustics 36
Directional pattern
  microphones 584
Directional sound reflection
  acoustics for auditoriums and concert halls 181
Directivity
  loudspeakers 638
  message repeaters and evacuation systems 1524
Directivity balloon curves
  computer aided sound system design 1343
Directivity data
  loudspeakers 632
Directivity deviation ratio
  computer aided sound system design 1348
Directivity factor
  computer aided sound system design 1346, 1349
  units of measurement 1660
Directivity index
  computer aided sound system design 1349
Direct-to-reverberant ratio
  psychoacoustics 61
DirectX
  consoles 985
Discrete absorbers
  acoustical treatment for indoor areas 108
Discrete echoes
  designing for speech intelligibility 1394
Discrete logic
  consoles 942
Discrete operational amplifiers
  consoles 838
Discrete solid-state devices 317
Discrete-time signals
  DSP technology 1162
Discs
  analog discs stereophonic groove 1015
  analog record 1015
  analog record dimensions 1015
  LP records 1015
Disk Description Protocol
  optical disc formats 1134
Dispersion
  fiber optics 480
Dispersion of light
  fiber optics 457
Displacement
  units of measurement 1651
Displacement of Maximum
  acoustics for auditoriums and concert halls 181
Display color
  display technologies 1580
Display technologies
  analog display 1581
  aspect ratio 1581
  Birefringence 1587
  Brilliant Color 1586
  candelas per meter squared 1579
  chrominance 1582
  CIE Chromaticity Diagram (Color Chart) 1580
  component video 1582
  composite video signal 1582
  contrast 1579
  Digital Content Protection 1582
  digital display 1581
  digital fixed matrix display scan conversion 1582
  Digital Light Processing (DLP) 1585
  Digital Micromirror Device (DMD) 1585
  digital video interface (DVI) 1582
  display color 1580
  DLP 1585
  DLP (Digital Light Processing) 1580
  DLP color wheel 1586
  DMD 1585
  fill factor 1580
  fixed matrix display 1582
  Flexible OLED 1588
  FOLED 1588
  Foot-Lamberts 1579
  HDMI 1583
  HDTV 1582
  High Definition Multimedia Interface (HDMI) 1583
  interlaced scanning 1581
  LCD 1584
  LCD projection TV 1585
  LCoS 1587
  LCoS projection system 1587
  LED 1588
  Light Emitting Diode 1588
  Liquid Crystal Display (LCD) 1584
  Liquid Crystal on Silicon (LCoS) 1587
  lumen 1579
  luminance 1582
  nits 1579
  NTSC television 1581
  OLED 1587
  Organic Light Emitting Diode (OLED) 1587
  PDP 1583
  plasma display 1583
  plasma monitor 1583

  progressive scan 1581
  resolution 1580, 1589
  scaling 1581
  scan conversion 1581
  scanning 1580
  single chip DLP technology 1586
  S-Video 1582
  TFT-LCD technology 1584
  three chip DLP projection TV 1586
  TOLED 1588
  Transparent OLED 1588
  video graphics array (VGA) 1582
  YPbPr 1582
DisplayPort
  wire and cable 411
Dissipation coefficient
  acoustics for auditoriums and concert halls 183
Dissipation factor (DF)
  capacitors 255
Distance cues
  psychoacoustics 60
Distance perception
  Psychoacoustics 60
Distortion
  attenuators 767, 779
  audio transformers 279, 283, 289
  delay 813
  fiber optics 469
  gain structure 1226
  hysteresis
    audio transformers 306
  IM
    audio transformers 284
  intermodal
    fiber optics 481
  loudspeakers 629
  multimode
    fiber optics 482
  phase 285
  virtual systems 1442–1443
  VU meters 998
distortion 1549
Distributed ceiling loudspeaker system
  designing for speech intelligibility 1395
Distributed ceiling loudspeaker systems
  sound system design 1269
Distributed column system
  sound system design 1270
Distributed line systems
  intercoms 1561
Distributed loudspeaker system
  sound system design 1268
Distributed loudspeaker systems
  designing for speech intelligibility 1398
  outdoor sound 203
Distributed mixing
  personal monitors 1430
Distributed systems outdoors
  sound system design 1323
Diversity antenna
  microphones 571
Dividing networks
  attenuators 772
DLP
  display technologies 1585
DLP (Digital Light Processing)
  display technologies 1580
DLP color wheel
  display technologies 1586
DMD
  display technologies 1585
Dolby Digital
  optical disc formats 1149
  surround sooound 1595
Dolby Digital (AC-3)
  optical disc formats 1148
Dolby Digital 5.1
  surround sound 1597
Dolby Digital Surround EX
  surround sound 1595
Dolby Pro Logic II
  surround sound 1598
Dolby Stereo
  surround sound 1593
Domain chart
  test and measurement 1613
Dome radiator
  loudspeakers 610
Dome tweeter
  loudspeakers 626
Doping agent
  solid state devices 318
Double insulated
  grounding and interfacing 1192
Double-diffused epitaxial mesa transistor
  solid state devices 326
Doublet source directivity
  loudspeaker clusters 658
DRAM
  message repeaters and evacuation systems 1528
Draperies
  acoustical treatment for indoor areas 108
Drift field transistor
  solid state devices 325
Drum machine
  MIDI 1116
Drum recording
  consoles 826
Drum throne shakers
  personal monitors 1430
Dry reed relays 381–382
Dry reed switches
  relays 387
DSD
  optical disc formats 1144
DSP
  delay 811
  equalization 963
  filtering 963
  optical disc formats 1140
  virtual systems 1440, 1443, 1447, 1451–1453
DSP arithmetic
  DSP technology 1173
DSP equalizers
  sound system design 1287
DSP filtering
  consoles 963
DSP technology
  algorithm development 1172
  analog to digital converter 1167
  bounded input/bounded output stable 1164
  causal system 1164
  code development 1176
  computational power 1176
  continuous time signal 1161
  continuous to discrete conver-

sion 1168
continuous-time signal 1167, 1169
convolution 1165
cosine sequence 1163
cosine signal 1162
data buffers 1174
delayed impulses 1164
digital signal processing 1161
digital signal processor 1172
discrete-time signals 1162
DSP arithmetic 1173
dynamic range 1173
finite impulse response linear filters 1167
FIR 1167
floating-point numbers 1173
Fourier analysis 1164
Fourier transform 1166, 1168
frequency domain 1165
I/O system 1177
IIR 1167
impulse response 1164
infinite impulse response linear filter 1167
linear time invariant systems 1164
loudspeakers 621
lowpass filter 1166
lowpass filtering 1162
MFLOPS 1176
microprocessor 1172
MP3 file 1161
MP3 players 1161
MPEG audio compression 1163, 1166
Nyquist Sampling Theorem 1169
programmable digital signal processor 1172
quantization 1170
quantization error 1171
real-time programming 1174
sampling rate 1171
sequences 1162
Sigma-Delta converter 1171
signal to noise ratio 1171
SNR 1171
stability 1164
system delay 1174
time invariance 1164
time-domain 1165
VLSI 1173
Z-transform 1166
DSP-driven line arrays
  loudspeaker clusters 662
DSP-driven vertical arrays
  loudspeaker clusters 657
DTMF
  intercoms 1563, 1575
Dual balanced line receiver
  integrated circuits 355
Dual channel FFT measurement
  loudspeakers 640
Dual driver earphones
  personal monitors 1424
Dual mono mixes
  personal monitors 1422
Dual Pro Pop pop screen
  microphones 578
Dual-channel FFT analyzers
  test and measurement 1616
Dual-coil latching relay 380
Dummy head
  computer aided sound system design 1355
Duplex
  audio transformers 300
dv/dt transients
  relays 389
DVD 1439
  optical disc formats 1133
DVD disc manufacturing
  Optical disc formats 1146
DVD family
  optical disc formats 1145
DVD file format
  optical disc formats 1146
DVD optical disc formats
  optical disc formats 1145
DVD+RW
  optical disc formats 1152
DVD-Audio
  optical disc formats 1133, 1145, 1148
DVD-R
  optical disc formats 1151
DVD-RAM
  optical disc formats 1152
DVD-ROM
  optical disc formats 1133, 1145, 1151
DVD-RW
  optical disc formats 1152
DVD-Video
  optical disc formats 1133, 1145, 1147, 1151
DVD-Video formats
  optical disc formats 1151
DVD-Video VAN disc
  optical disc formats 1147
Dynamic changes
  psychoacoustics 54
Dynamic earphones
  Personal monitors 1424
Dynamic line regulation
  power supplies 672
Dynamic load regulation
  power supplies 672
Dynamic microphone 503
Dynamic range
  delay 810
  DSP technology 1173
  gain structure 1225–1227
  psychoacoustics 51
  virtual systems 1442
Dynamics processing
  consoles 895
Dynamics processors
  virtual systems 1439

E

E. Norman Rauland 12
Ear
  acoustic reflex 45
  anatomy and function 43
  auditory nerve 1433
  auditory neurons 47
  auditory system 45
  auricle 43
  basilar membrane 46
  bone conduction 45
  buckling motion 45
  canal 45
  characteristic frequency 46
  cilia 1433
  closed pipe 45
  cochlea 43, 46, 1433
  concha 43, 45
  cross-section view 43
  cross-sectional sketch of the co-

chlea 46
ear canal 43, 1432
ear drum 43, 1432
endolymph 46
eustachian tube 43, 45
Fourier analyzer 51
frequency selectivity 47
hair cells 43
harmonic distortion 50
hearing damage 45
helicotrema 46
incus 43, 45
inner 43, 45
inner hair cells (IHC) 46
labyrinth 45
lever motion 45
malleus 43, 45
masking 48
mastoid 44
mechanical system of the middle ear 46
middle 45
middle ear 1432
middle ear infection 45
neural firings 46
nonlinearity 49
Organ of Corti, 46
ossicles 45, 1432
ossicular chain 45
otoacoustic emissions 50
ototoxic medicines 47
outer 43
outer ear 1432
outer hair cells (OHC) 46
oval window 43, 46
perilymph 46
personal monitors 1432
pharynx 43
phase locking 50
pinna 43, 1432
piston motion 45
Reissner's membrane 46
round window 43, 46
scala media 46
scala tympani 46
scala vestibuli 46
squamous 44
stapedius muscle 45
stapes 43, 45, 1433
stereociliae 46
tectorial membrane 46
temporal bone 44
tensor tympani 45
traveling wave 47
tuning curve 46
tuning curves 47
tympanic 44
tympanic membrane 43
vestibular system 45
vibration patterns 46
wavelet analysis 51

Ear canal
  personal monitors 1432
Ear canal (auditory meatus)
  psychoacoustics 43
Ear drum
  personal monitors 1432
Ear external 43
Ear middle 43
Ear molds
  psychoacoustics 59
Ear-buds
  personal monitors 1423, 1433
Eardrum
  fundamentals of audio and acoustics 34
  psychoacoustics 43
Early reflections
  designing for speech intelligibility 1388, 1394
Early to late ratios
  designing for speech intelligibility 1404
Early-decay time (EDT)
  test and measurement 1627
Earmuffs
  what's the ear for? 1635
Earphone transducers
  personal monitors 1424
Earphones
  personal monitors 1422
Earplugs
  what's the ear for? 1635
Earth ground
  consoles 846
  grounding and interfacing 1190
EASE 648, 651
EASE SpeakerLab
  computer aided sound system design 1368
EASERA 1626
EASERA SysTune
  computer aided sound system design 1381
EBU 1443
EBU standard
  VI meters 1005
Echo
  consoles 828, 962
  sound system design 1245
Echo canceller
  preamplifiers and mixers 760
Echo criterion
  acoustics for auditoriums and concert halls 154
Echo disturbance
  designing for speech intelligibility 1398
Echoes
  designing for speech intelligibility 1397
Echoes outdoors
  Sound system design 1323
Echogram
  acoustical modeling and auralization 221
EDC/ECC
  Optical disc formats 1141
Eddy current loss
  eddy current lossg 1058
Eddy currents 1058
  audio transformers 280
Eddy-current losses
  audio transformers 279
Edge emitting LED
  fiber optics 463
Editing
  MIDI 1120
EDT
  test and measurement 1627
Edward C. Wente 9
Edward Wente 8
Edwin Howard Armstrong 16
Edwin Kennelly 6
Edwin W. Rice 6
EEPROM
  message repeaters and evacuation systems 1528
Effect of directional loudspeakers
  sound system design 1253

Effect of directional microphones
  sound system design 1243, 1253
Effect of humidity
  sound system design 1244
Effect of temperature layers
  sound system design 1244
Effect of wind
  Sound system design 1326
  sound system design 1244
Effective tonearm length
  analog discs 1023
Effects of various loads
  relays 379
Effects processing
  personal monitors 1431
Effects send
  consoles 828
Efficiency
  computer aided sound system design 1347
  loudspeakers 609, 627
EFM encoding
  optical disc formats 1137
EFM signal
  optical disc formats 1140
EFMPlus modulation
  optical disc formats 1147
Eigenfrequencies
  acoustics for auditoriums and concert halls 192
Electret microphone 515
Electric capacitance (C)
  units of measurement 1650
Electric charge (Q)
  units of measurement 1650
Electric conductance (G)
  units of measurement 1650
Electric inductance
  units of measurement 1651
Electric potential difference (V)
  units of measurement 1650
Electric resistance (R)
  units of measurement 1650
Electrical loading
  analog discs 1033
  cartridge 1033
Electrical power required (EPR)
  sound system design 1241
  units of measurement 1660
Electrical reflections
  fundamentals of audio and acoustics 29
Electrical test instruments
  test and measurement 1607
Electrical testing
  test and measurement 1607
Electro-acoustical transducers
  computer aided sound system design 1342
Electrodynamic cartridges
  analog discs 1027
Electrodynamic transducers
  loudspeakers 598
Electromagnetic interference (emi)
  power supplies 682
  relays 379
Electromagnetic spectrum
  fiber optics 453
Electromagnetic wave
  fundamentals of audio and acoustics 27–28
Electromagnetic waves
  grounding and interfacing 1181, 1184
Electromechanical relays 379
Electromotive force
  fundamentals of audio and acoustics 32
Electronic architecture
  acoustics for auditoriums and concert halls 190–191
Electronic balanced outputs
  consoles 869
Electronic crossovers
  sound system design 1288
Electronic Industries Association (EIA) output sensitivity
  microphones 517
Electronic switching
  consoles 851
Electronic system
  assistive listening systems 1549
Electronically balanced input
  audio transformers 302
Electrostatic fields
  grounding and interfacing 1187
Electrostatic loudspeaker 603
Electrostatic loudspeakers 604
Electrostatic shield
  audio transformers 282
Electrostatic transducers
  loudspeakers 602
Elliptical effect
  acoustics for auditoriums and concert halls 180
Elliptical filter
  filters and equalizers 793
Elliptical stylus
  analog discs 1031
Emergency call systems
  intercoms 1561
Emergency crisis communications
  intercoms 1566
EMI
  digital audio interfacing and networks 1503
Empirical mass law
  acoustical noise control 71
Empirical models 229
  acoustical modeling and auralization 215
Enclosure spring constant
  loudspeakers 635
Energy (E)
  units of measurement 1649
Energy criteria
  acoustics for auditoriums and concert halls 152
Energy Time Curve
  delay 808
  small room acoustics 136
Enkasonic floor system
  acoustical noise control 80
Envelope Time Curve
  small room acoustics 136
Envelope-Time Curve
  test and measurement 1618
Envelope-Time Curve (ETC)
  Test and measurement 1616
  test and measurement 1612
Envelopment
  delay 808
Environmental temperature effects
  relays 382
EPA
  what's the ear for? 1633
EPROM
  message repeaters and evacua-

tion systems 1528
Equal loudness contours
  psychoacoustics 51
Equal temperaments
  psychoacoustics 55
Equalization
  consoles 875
  designing for speech intelligibility 1389, 1400
  sound system design 1307
  virtual systems 1439
Equalization control
  consoles 890
Equalization test setup
  sound system design 1309
Equalizer
  consoles 875, 894
  digital dynamics 970
  digital limiter 972
  high-frequency response anomalies 970
  literal dynamics 972
  mapping dynamics 971
  parametric equalizer
    virtual systems 1449
  shelving 969
  shelving equalizer 969
Equalizer response curves
  consoles 876
Equalizers
  filters and equalizers 785, 800
  microphones 581
  sound system design 1286
Equalizing a distributed ceiling system
  sound system design 1273
Equalizing batteries
  power supplies 694
Equal-level far-end crosstalk (EL-FEXT)
  wire and cable 415
Equal-loudness contour curves
  surround soound 1599
Equal-loudness contours
  fundamentals of audio and acoustics 35
Equations for finding surface areas 1680
Equipment locator
  intercoms 1568
Equivalent acoustic distance
  units of measurement 1658
Equivalent acoustic distance (EAD)
  sound system design 1242
Equivalent circuits
  transistors 328
Equivalent input noise (EIN)
  preamplifiers and mixers 738
  units of measurement 1661
Equivalent rectangular bandwidth (ERB)
  psychoacoustics 49
Equivalent series inductance (ESL) 255
Equivalent series resistance (ESR) 255
Equivalent sound absorption area
  acoustics for auditoriums and concert halls 165
Erase heads
  erase headscording 1063
ERB number
  psychoacoustics 49
ERES (Electronic Reflected Energy System)
  Acoustics for auditoriums and concert halls 191
Error concealment algorithms
  Optical disc formats 1140
ETC
  Test and measurement 1616
Ethernet
  digital audio interfacing and networks 1487
  equipment 1490
  switches 1492
  virtual systems 1448
  wire and cable 412
Ethernet (Cat-5)
  Personal monitors 1430
Ethernet packet format
  digital audio interfacing and networks 1488
Ethernet Star of Stars topology
  digital audio interfacing and networks 1491
Ethernet style connectors
  sound system design 1296
EtherSound
  digital audio interfacing and networks 1510
Euler 6
Euro-Block connectors
  sound system design 1293
Eustachian tube
  psychoacoustics 43
Evaluating loudspeaker sound quality
  sound system design 1259
Examples of English Words Used in Intelligibility Tests 157
Exceedance level
  acoustical noise control 70
Excess loss
  fiber optics 475
Excitation current
  audio transformers 276–277, 280, 283
Excitation flux
  Audio transformers 277
Expandor
  consoles 906
Expansion
  consoles 907
Explosion proof drivers
  loudspeakers 602
External ear
  psychoacoustics 43
Externalization
  psychoacoustics 59
Extruded heat sinks 369

F

Factory floor applications
  message repeaters and evacuation systems 1517
Fader automation
  consoles 936
Faders
  consoles 946
Fairchild µA741
  integrated circuits 340
Fake inductor
  consoles 883
Fans
  heatsinks 369
Faraday shield
  audio transformers 282, 291–292
  consoles 863

  grounding and interfacing 1212
Far-end crosstalk (FEXT)
  wire and cable 415
Fast Ethernet
  digital audio interfacing and networks 1490
Fast Fourier Transform
  test and measurement 1615
Fast link pulses (FLPs)
  digital audio interfacing and networks) 1494
FC/PC connector 471
FDDI fiber optic connector 472
Federal Communications Commission (FCC)
  Assistive listening systemsAssistive listening systems
    Federal Communication Commision (FCC) 1550
Feed forward VCA-style dynamics
  consoles 907
Feedback
  personal monitors 1429
  sound system design 1241, 1264
Feedback factor
  audio transformers 287
Feedback phase-leading stability compensation
  consoles 839
Feedback stability margin
  units of measurement 1659
Feedback stability margin (FSM)
  sound system design 1241
Feedback theory 707
Feedback-style compressors
  consoles 973
Feedback-style limiting
  consoles 973
Feedback-type volume control 778
Ferrite beads 268
  grounding and interfacing 1210
Ferrofluids
  loudspeakers 608
Ferro-magnetic materials
  audio transformers 278
Ferrule
  fiber optics 480
FFT measurements
  test and measurement 1615
Fiber Data Distributed Interface (FDDI)
  fiber optics 480
Fiber optic 480
  connector loss 460
  detector specifications 476
  multimode step index fiber 456
  Optical Time Domain Reflectometer (OTDR) 462
  refraction 454
Fiber optic cable
  digital audio interfacing and networks 1502
Fiber optic cable bend radius
  digital audio interfacing and networks 1503
Fiber optic connectors
  digital audio interfacing and networks 1502
Fiber optic link
  fiber optics 470, 480
Fiber optic systems 1572
Fiber optic transmission loss 458
Fiber optics 451
  absorption 459
  advantages of using fiber optics for audio 452
  Alexander Graham Bell 451
  angle of incidence 454–455
  angle of reflection 455
  angle of refraction 454
  attenuation 458, 477, 479
  attenuation coefficient 458
  attenuation in fiber 459
  attenuation measurement 461
  avalanche photodiode (APD) 467
  bandwidth 468
  beamwidth 479
  bend radius fibers 459
  biconic connector 471
  bit-error-rate 470
  Burrus LED 463, 480
  cable specifications 476
  characteristics of cables 457
  coherent light 465
  connector installation 473
  connectors 1502
    couplers 475
    D4 472
    FC/PC 471
    FDDI 472
    LC 472
    LC type 474
    MT-RJ 473
    SC 472
    SFF 472
    SMA 471
    ST 472
  core 451
  coupling loss 460
  CSL LightSplice System 474
  cut-back method test 461
  dense wavelength division multiplexing 480
  design procedures 477
  detector specifications 476
  detectors 467
  dispersion of light 457
  distortion 469
  edge emitting LEDs 463
  electromagnetic spectrum 453
  Fiber Data Distributed Interface (FDDI) 480
  fiber optic system design 475
  fiber optic transmission loss (FOTL) 458
  fiber specifications 476
  Fibrlok splice 475
  Fresnel reflection 454, 480
  gigabit optical transceivers 469
  history 451
  index of refraction 481
  insertion loss 471–472, 475, 481
  integrated detector-preamplifier (IDP) 467
  intermodal distortion 481
  laser diode (LD) 465, 481
  LED 463
    lifetime 467
    speed 467
  light 452, 481
  light emitting diode (LED) 481
  light prism 453
  light propagation 456
  light rays 454
  light source specifications 476
  link 480
  loose tube fiber jackets 457

microbend loss 459
microlensed LED 463
monochromatic light 465, 481
multimode distortion 482
multimode fiber 482
multimode fiber optic cable 1502
multimode graded index fiber 456
noise 468
numerical aperture 460, 482
optical fiber class 482
optical fiber loss 458
Optical Time Domain Reflectometer 462
Optical Time Domain Reflectometer (OTDR) 482
optoelectronic 482
optoelectronic integrated circuits (OEIC) 482
OTDR 462
particle of light 453
passive optical interconnections 470
performance 456
photodiode 482
Photophone 451
PIN photodiode 467, 482
PIN-FET receiver 482
Plank's Law 453
plastic optical fiber (POF) 482
population inversion 465
positive-intrinsic-negative detector PIN 454
quantum efficiency 468
Rayleigh scattering 459, 483
receiver 483
  sensitivity 483
  specifications 476
reflection 454, 483
refraction 454, 483
refractive index 455, 483
repeaters 469, 483
scattering 459
signal loss 458
single mode fiber 456, 483
single optical fiber 454
single-mode fiber cable 1502
Snell's Law 455
splice 474–475, 477, 483
splice loss 460
spontaneous emission 483
ST connector 483
substrate 481
superluminescent diodes (SLD) 466, 483
surface emitting LED 463
system specifications 475
thermal noise 468
tight buffer fiber jackets 457
transceivers 469
transmission Loss 483
transmitter 483
  specifications 476
transmitter/receiver modules 469
types of fiber 456
UniCam Connector System tool 474
vertical cavity surface emitting laser (VCSEL) 466, 483
wavelength 452, 464, 466, 482
wavelength division multiplexing (WDM) 484

Fiber specifications
  fiber optics 476
Fibrlok splice
  fiber optics 475
Field distortion
  assistive listening systems 1548
Field effect transistor
  transistors 326
Field effect transistors
  amplifier 718
Field Programmable Gate Array (FPGA)
  consoles 976
Field strength
  assistive listening systems 1547
Field-effect transistor attenuator 779
Field-effect transistors
  consoles 851
Filament
  tubes 311
Fill factor
  display technologies 1580
Filter
  active filter 710–711
  EMI filter 724
Filter gain
  consoles 888
Filter quantization distortion
  consoles 967
Filter slopes
  loudspeakers 623
Filtering
  comb 1432
Filters
  bandpass 884
  Butterworth
    virtual systems 1449
  filter gain 888
  FIR 963, 1454
  IIR filters 966
  interpolation 958
  power supplies 674
  reconstruction 958
  transversal 963
Filters and equalizers
  active 793
  adaptive 802
  all pass 796
  arithmetic symmetry 785
  band pass 785, 791
  band-reject 791
  bandwidth 785
  biquad 799
  biquad IIR 799
  Butterworth 792
  capacitive networks 787
  center frequency 785
  Chebyshevs 792
  constant Q 802
  constant-k 789
  corner frequency 785
  cutoff frequency 785
  damping factor 793
  digital 797
  elliptical 793
  equalizers 785, 800
  finite impulse response 798
  FIR 798
  Fourier Transform 798
  geometric symmetry 785
  graphic 800
  group delay 786

  high pass 785, 789
  Hilbert Transform 786
  impedance matching 793
  impulse 786
  inductive networks 788
  ladder networks 791
  linear phase 798
  linear time invariant 786
  Linkwitz-Riley 792
  low pass 785, 789
  minimum phase 786
  normalizing 793
  one-third octave 800
  parallel resonant circuit 790
  parametric 801
  pass band 785
  passive filters 785–786
  phase angle 786
  phase delay 786
  pole-zero analysis 796
  potential divider 787
  programmable 802
  Sallen-Key 794
  scaling 793
  second order 788
  semi-parametric 801
  series resonant circuit 790
  state variable 795
  stop band 785
  swept frequency 801
  switched capacitor 797
  T network bandpass 791
  transient response 786
  transition band 785
  transversal 801
  unity gain Sallen-Key 795
Finite Element Method (FEM)
  computer aided sound system design 1354
Finite impulse response
  optical disc formats 1140
Finite impulse response (FIR) linear filters
  DSP technology 1167
Finite impulse response filter
  filters and equalizers 798
Finite-difference time-domain
  acoustical modeling and auralization 228
FIR
  DSP technology 1167
  optical disc formats 1140
FIR coefficient sets
  consoles 967
FIR filter
  filters and equalizers 798
FIR filters
  consoles 963
  virtual systems 1454
Fire safety codes
  acoustical treatment for indoor areas 121
Firewire 1487
FireWire serial digital
  wire and cable 410
Fitzroy
  units of measurement 1661
Fixed matrix display
  display technologies 1582
Fixed threshold automatic microphone mixers 748
Flange
  delay 809
Flash conversion
  consoles 956
Fleming valve
  tubes 311
Fletcher-Munson curves
  psychoacoustics 51
Flexible OLED
  display technologies 1588
Floating floors
  acoustical noise control 79
Floating point processors 960
  consoles 960
Floating-point numbers
  DSP technology 1173
Floor and ceiling systems
  acoustical noise control 79
Floor construction
  acoustical noise control 78–79
Floppy disks
  magnetic recording 1083
Flow resistamce
  acoustics for auditoriums and concert halls 184
Flutter
  flutter 1047
  flutterc recording 1091
Flutter echo
  delay 808
Flux
  audio transformers 275
  fluxetic recording 1067
flux 275
Flux density 266
  audio transformers 275, 278–279, 285
  flux densityording 1067
  grounding and interfacing 1189
  loudspeakers 636
Flux lines
  flux linesecording 1067
Flux loop
  flux looprecording 1094
FM Broadcast
  assistive listening systems 1550
FM interpretation systems
  interpretation and tour group systems 1534
FM receiver
  interpretation and tour group systems 1540
FM transmission
  intercoms 1575
FM transmitter
  interpretation and tour group systems 1539
Focused reflections
  fundamentals of audio and acoustics 31
FOH
  personal monitors 1418, 1427
Foil shields
  wire and cable 426
Foil-braid-foil-braid cable
  wire and cable 427
Foldback
  consoles 828
Folded fin heatsinks 369
Folded horns
  loudspeakers 614
FOLED
  display technologies 1588
Foot Lamberts
  display technologies 1579
Force (F)
  units of measurement 1650
Forced convection

heatsinks 367
Forced-convection cooling
heatsinks 373
Forward masking
psychoacoustics 48
Forward resistance
power supplies 669
Forward voltage drop
power supplies 669
Forward-current-transfer ratio
transistors 329
Fourier analysis
DSP technology 1164
Fourier analyzer
psychoacoustics 51
Fourier Transform
filters and equalizers 798
Fourier transform
DSP technology 1166, 1168
loudspeakers 640
test and measurement 1613
Frame check sequence (FCS)
digital audio interfacing and networks 1489
Frank Jewett 7
Frank McIntosh 5
Franssen effect
psychoacoustics 60
Frederick Vinton Hunt 15
Free-field open-circuit sensitivity
computer aided sound system design 1356
Free-field sensitivity
computer aided sound system design 1344
Freestanding radiator
loudspeakers 635
French braid shields
wire and cable 426
French Système International d'Unités
units of measurement 1647
Frequency
fundamentals of audio and acoustics 26
Frequency (f)
units of measurement 1651
Frequency difference limen
psychoacoustics 56
Frequency domain
DSP technology 1165
Frequency range
psychoacoustics 51
Frequency resolution
loudspeakers 639
Frequency response 1388
audio transformers 285
computer aided sound system design 1345
fundamentals of audio and acoustics 27
loudspeakers 628
microphones 498
test and measurement 1613
Frequency response of a cardioid microphone 497
Frequency selectivity
psychoacoustics 52
Fresnel reflection
fiber optics 454, 480
Fringing
fringing recording 1057
Front to random factor
computer aided sound system design 1346
Front to random index
computer aided sound system design 1346
Front-end instability
consoles 842
FSTC 72
Full messages
MIDI 1129
Full wave bridge rectifier
power supplies 673
Full wave power supplies 672
Function modes
consoles 923
Fundamental mode
fiber optics 480
Fundamental quantities in physics 1647
Fundamentals of audio and acoustics
absolute levels 25
absorption 29
acoustic gain 32
acoustic impedance 28
acoustic pressure levels 26
acoustic reflections 28
acoustic sources 36
acoustic wave 28
acoustic wavelengths 30–31
ampere 32
antennas 28
architectural acoustics 28
audible spectrum 27
audio program material 35
Bel 24
capacitive reactance 33
complex impedance 33
constant voltage 25
convex surface 29
cosine wave 30
crest factor 36
decibel conversion 24
decibel formulas 26
decibel references 25
directional energy radiation 36
eardrum 34
electrical reflections 29
electromagnetic wave 27–28
electromotive force 32
equal-loudness contours 35
focused reflections 31
frequency 26
frequency response 27
harmonic motion 31
human auditory system 34
human hearing 26, 34
image shift 31
impedance 25, 33
inductive reactance 33
inverse-square law 37
level 26
level difference 26
line arrays 38
line source 37
logarithmic scale 23
loudness 26
loudspeakers 36
multiple loudspeakers 31
nominal impedance 33
ohm 32
Ohm's Law 32
open circuit 25
peak loudness 36
phase interference 32
phase shift 33
point source 36

- power 25
- power amplifier 25
- program material 34
- radio frequencies 28
- rate of electron flow 32
- rated impedance 33
- reactance 33
- reflected wave 29
- relative loudness 36
- relative phase 31
- signal chain 25
- sine wave 30
- sound propagation 36
- spectral response 27
- speed of light 27
- superposition 30
- surface shapes 29
- time period 27
- time response 34
- Tinnitus 35
- tonal coloration 31
- typical sound pressure levels 26
- VI meter 36
- volt 32
- wavelength 27
- weighting curves 34
- weighting scales 35

Fuses
- sound system design 1278

G

G. S. Ohm 6

G.R. Kirchhoff 6

Gain and bandwidth
- integrated circuits 338

Gain and headroom
- preamplifiers and mixers 737

Gain sharing automatic microphone mixers 748

Gain structure
- A/D conversion 1230
- amplification factor 1227
- amplifier power ratings 1231
- amplifier sensitivity 1230
- audio waveform fundamentals 1224
- average voltage 1224
- bridged interfaces 1224
- clipping 1225
- constant current interface 1224
- constant voltage 1223
- constant voltage interface 1223
- continuous output power 1231
- crest factor 1224–1225, 1231
- distortion 1226
- dynamic range 1225–1227
- interfaces 1223
- loudspeaker power ratings 1231
- matched interface 1223
- optimized method 1229
- peak amplitude 1225
- peak to peak voltage 1224
- peak voltage 1224
- power ratings 1231
- power transfer 1228
- program material 1231
- resistive pads 1229
- root-mean-square voltage (rms) 1224
- signal-to-noise ratio 1227
- sine wave power 1231
- sinusoidal waveform 1227
- SNR 1226–1227
- thermal noise 1225–1226
- unity 1225
- unity amplification method 1228
- voltage amplitude 1227
- VU meter 1228

Gain-bandwidth product
- transistors 329

Gain-before-feedback
- Personal monitors 1417

Galleries
- acoustics for auditoriums and concert halls 173

Gapping
- audio transformers 279

Gate attack
- consoles 905

Gated-noise test
- test and measurement 1611

Gating
- consoles 905, 973

Gating feed-forward sidechain
- consoles 905

Gel/Cell batteries 696

General concert halls
- acoustics for auditoriums and concert halls 164

Generator impedance
- audio transformers 285

Generic loudspeaker library (GLL)
- computer aided sound system design 1362

Geometric symetry
- filters and equalizers 785

Geometrical models
- acoustical modeling and auralization 220

Geometrical reflection
- acoustics for auditoriums and concert halls 178

Geometrically complex room
- sound system design 1253

George A. Campbel 8

Germanium
- solid state devices 317

Germanium crystals
- solid state devices 319

Gigabit Ethernet
- digital audio interfacing and networks 1490

Gigabit optical transceivers
- fiber optics 469

Gilbert Cell
- consoles 937

Glass fiber panel
- acoustical treatment for indoor areas 103

Glossary of terms
- relays 374

Gordon Gow 5

Graded index fiber
- fiber optics 480

Graphic equalizer
- filters and equalizers 800

Graphic equalizers
- sound system design 1286

Greek alphabet 1663

Grid bias
- tubes 314

Grid current
- tubes 315

Grid voltage
- tubes 315

Grid-to-plate capacitance
- tubes 314

Ground current summing

consoles 848
Ground dummy
  grounding and interfacing 1204
Ground fault interrupters
  wire and cable 442
Ground isolators
  Grounding and interfacing 1213
  grounding and interfacing 1207
Ground lifters
  grounding and interfacing 1191
Ground loop
  grounding and interfacing 1193
Ground loops
  sound system design 1304
  wire and cable 428
Ground noise
  audio transformers 282
  consoles 849
Ground plane
  grounding and interfacing 1192, 1215
Ground-fault circuit interrupter (GFCI)
  grounding and interfacing 1215
Grounding
  consoles 844, 913
  grounding and interfacing 1181, 1189
  microphones 520
  sound system design 1301
Grounding and interfacing
  ac circuits 1182–1183
  back emf 1183
  balanced cable 1200
  balanced input 1208
  balanced input circuits 1213
  balanced interfaces 1212
  balanced output 1208
  balanced power 1215
  bandwidth 1203
  basic electronics 1181
  cable 1184
  cable capacitance 1187
  cable shield grounding 1213
  capacitance 1182
  capacitive coupling 1187
  capacitive reactance 1182
  CATV 1208
  characteristic impedance 1185–1186
  circuit theory 1181
  CMRR 1207, 1212
  common impedance coupling 1202
  common-mode rejection ratio (CMRR) 1185, 1196
  common-mode voltage 1185
  differential amplifier 1185
  double insulated 1192
  earth ground 1190
  electromagnetic waves 1181, 1184
  electrostatic fields 1187
  Faraday shield 1212
  ferrite beads 1210
  flus density 1189
  ground dummy 1204
  ground fault circuit interrupter (GFCI) 1215
  ground isolators 1207, 1213
  ground lifters 1191
  ground loop 1193
  ground plane 1192, 1215
  grounding 1181, 1189
  IEC tests for CMRR 1199
  impedance 1182–1183
  impedance matching 1186
  inductance 1182
  inductive coupling 1188
  inductor 1183
  interfaces 1185
  inter-outlet ground voltage noise 1194
  isolated grounding 1214
  law of induction 1183
  leakage current noise 1194
  line driver 1187
  magnetic field 1183–1185, 1188
  multipoint grounding 1192
  Ohm's Law 1181
  permeability 1189
  power line noise 1193
  power transformers 1190, 1194
  propagation velocity 1184
  RF interference 1203, 1210
  safety grounding 1190–1191
  shield-current induced noise 1201
  shielding 1187–1188
  signal grounding 1192
  signal quality 1212
  slew rate 1187
  standing wave ratio (SWR) 1185
  standing waves 1185
  static fields 1181
  surge protection 1216
  symmetrical power 1215
  system grounding 1181
  system noise 1193
  transient response 1212
  transmission lines 1184
  unbalanced cable 1202
  unbalanced interface 1213
  voltage divider 1186
  voltage matching 1186
  wire 1183
Grounding for safety outdoors
  sound system design 1303
Grounding to reduce external noise pickup
  sound system design 1304
Grounds of shields
  wire and cable 427
Group delay
  filters and equalizers 786
Group delay distortion 711
Grouping
  consoles 831
Grown-junction transistor
  transistors 325
Guitars
  MIDI 1118
Gypsum cavity wall absorption
  acoustical modeling and auralization 229
Gypsum wallboard
  acoustical noise control 74
Gypsum wallboard (GWB)
  acoustical treatment for indoor areas 111
Gyrator
  consoles 880
Gyrator types
  consoles 880

H

H pad 773

H. D. Arnold 8
H.E.A.R. 1435
H.H. Scott 5
Haas Effect
  personal monitors 1431
Haas effect
  psychoacoustics 60
Hair cells
  psychoacoustics 43
Hamming
  Test and measurement 1623
Handheld entertainer microphone 530
Hang time
  consoles 903
Hann window
  test and measurement 1623
Hard disk drives
  magnetic recording 1082, 1084
Hard disk electronics
  magnetic recording 1086
Hard failure
  relays 375
Hard wired systems
  interpretation and tour group systems 1533
Harmonic distortion
  psychoacoustics 50
  VI meters 998
Harmonic motion
  fundamentals of audio and acoustics 31
Harold D. Arnold 8
Harold S. Black 10
Harry Nyquist 10
Harry Olson 13
Harvey Fletcher 9
HD DVD
  optical disc formats 1152
HDMI
  display technologies 1583
HD-SDI
  wire and cable 421
HDTV
  display technologies 1582
Head relapping 1095
Head related transfer function (HRTF)
  psychoacoustics 58
Head shadow
  psychoacoustics 58
Head worn microphone 528
Headphone diffuse-field equalizers
  microphones 554
Headphones
  binaural reproduction 232
Head-related transfer functions
  computer aided sound system design 1355
Head-related transfer functions (HRTF)
  acoustical modeling and auralization 232
Headroom
  consoles 933
  sound system design 1241, 1275
Headset
  infrared 1538
Headsets
  intercoms 1575
Hearing aid response
  assistive listening systems 1546
Hearing conservation
  personal monitors 1418
Hearing damage
  personal monitors 1433
  psychoacoustics 45, 52
  what’s the ear for? 1633, 1635
Hearing impaired
  assistive listening systems 1545
hearing impaired systems 1556
Hearing loss
  psychoacoustics 47
  What’s the ear for? 1633
Hearing tests
  what’s the ear for? 1635
Hearing threshold
  psychoacoustics 51
Heat conduction
  loudspeakers 608
Heat Sink Design 370
Heat transfer
  loudspeakers 609
Heat transfer mechanisms
  loudspeakers 609
Heater
  tubes 317
Heating, ventilating, and air conditioning (HVAC)
  acoustical noise control 89
Heatsink materials 370
Heatsinks
  air flow management 369
  aluminum extruded 368
  audio cooling tube 369
  bonded fin 369
  case temperature 373
  conduction 367
  convection 367, 370
  extruded 369
  fans 369
  folded fin 369
  forced convection 367
  forced-convection cooling 373
  materials 370
  minimum surface area 372
  natural convection 367
  plenum 369
  radiation 368, 370
  specific thermal resistance 372
  temperature differential 369
  temperature gradient 367
  thermal conductivity 370
  thermal performance 367
  thermal resistance 370–372
  thermal solutions 368–369
  thermal spreading resistance 368
  thermally conductive adhesives 371
  transistor mounting kit 371
Helical-scan digital tape transports
  magnetic recording 1080
Helmholtz resonators
  acoustical treatment for indoor areas 111
  acoustics for auditoriums and concert halls 185, 192
Hendrik Bode 10
Heptode
  tubes 311
Hermetic seal
  relays 375
Hewlett Packard 18
Hexode

  tubes 311
Hhuman ear
  sound system design 1260
High Definition DVD
  optical disc formats 1152
High Definition Multimedia Interface (HDMI)
  display technologies 1583
  wire and cable 410
High frequency beaming
  loudspeaker clusters 656
High frequency cable
  wire and cable 431
High pass filters
  filters and equalizers 785, 789
High peak-transient-voltage protection
  relays 389
High speed switching relays 386
High surge-current protection
  relays 389
High transient voltage
  relays 377
High-frequency attenuation in air
  sound system design 1322
High-frequency control
  consoles 894
High-frequency, high-date rate cable
  wire and cable 409
High-hysteresis materials
  audio transformers 279
High-loss frame walls
  acoustical noise control 74
High-pass filter 714
Highpass filter
  loudspeakers 622, 635
High-pass filters
  consoles 875, 888
  sound system design 1288
High-power regulated power supply 682
Hilbert Transform
  filters and equalizers 786
  test and measurement 1618
History
  fiber optics 451
Hold time
  consoles 903
Holophone surround sound microphone 551
Homemade acoustical doors
  acoustical noise control 83
Homemade acoustical windows
  acoustical noise control 86
Horizontal cabling
  wire and cable 412
Horizontal directivity
  loudspeaker clusters 663
Horizontal woofer arrays
  loudspeaker clusters 652
Horn directivity
  loudspeakers 638
Horn loudspeakers
  computer aided sound system design 1350
Horns
  loudspeakers 611
Hospital applications
  message repeaters and evacuation systems 1516
Hospital facilities intercoms 1566
Host micro controller
  consoles 981
Hot switching
  relays 375
Hot-spot temperature
  of resistors 244
House Ear Institute 1435
Houses of worship
  personal monitors 1429
Hub
  wire and cable 412
Hum
  audio transformers 296
Human auditory system
  fundamentals of audio and acoustics 34
Human ear
  computer aided sound system design 1355
Human external-ear transfer function
  microphones 554
Human hearing
  fundamentals of audio and acoustics 26, 34
Human hearing system
  test and measurement 1619
Human perception
  test and measurement 1625
Human voice
  computer aided sound system design 1343
Hybrid integrated circuits
  solid state devices 335
Hybrid models
  acoustical modeling and auralization 215, 230
Hydrophone
  psychoacoustics 45
HyMu®
  audio transformers 283
Hyperbolic effect
  acoustics for auditoriums and concert halls 180
Hyper-cardioid microphone 495
Hysteresis
  audio transformers 279, 284
  relays 375
Hysteresis distortion
  audio transformers 284, 306
Hysteresis loop
  audio transformers 278
Hysteresis loss
  hysteresis lossing 1058

I

I/O system
  DSP technology 1177
Identification of HVAC noise producers
  acoustical noise control 90
Idling noise
  delay 814
IEC 268-16 1407
IEC 60268-3
  audio transformers 302
IEC 76 1536
IEC 764 1531
IEC Standards 1665
IEC tests for CMRR
  grounding and interfacing 1199
IFFT
  Test and measurement 1615
IIR
  DSP technology 1167
IIR filters
  consoles 966
IM distortion
  audio transformers 284–285

Image modeling
  computer aided sound system design 1371
Image models
  acoustical modeling and auralization 220
Image shift
  fundamentals of audio and acoustics 31
Impact of absorption
  acoustical treatment for indoor areas 115
Impedance
  attenuator mismatch ratio 770
  characteristic impedance of VU meters 997
  fundamentals of audio and acoustics 25, 33
  grounding and interfacing 1182–1183
  inductors 265
  input-output of an amplifier 703
  Load
    audio transformers 297
  loudspeakers 629
  microphones 522
  nominal 33
  of a capacitor 258
  of an attenuator 770
  rated 33
  resistors, capacitors, inductors in circuits 268
  VU meters 998, 1000
Impedance (Z)
  relays 375
Impedance compatibility
  sound system design 1297
Impedance discontinuity
  relays 375
Impedance level correction
  VU meters 1000
Impedance matching
  attenuators 769
  filters and equalizers 793
  grounding and interfacing 1186
  sound system design 1297
Impedance tube testing methods
  acoustical treatment for indoor areas 100
Impulse
  filters and equalizers 786
Impulse response
  acoustical modeling and auralization 234
  consoles 963
  DSP technology 1164
  loudspeakers 624, 628
  test and measurement 1612–1613, 1618
Impulse response of a loudspeaker/room combination
  test and measurement 1620
Impulse response tests
  test and measurement 1628
Inadequate samples for limiters
  consoles 974
Incandescent Lamps 379
Incoherent
  fiber optics 480
Incus
  psychoacoustics 43
Index matching material
  fiber optics 481
Index of refraction
  fiber optics 481
Indoor sound reinforcement system
  sound system design 1245
Induced magnet cartridges
  analog discs 1028
Inductance
  coil 266
  consoles 879
  grounding and interfacing 1182
  in inductors 263
  magnetic field 1546
  units of measurement 1651
Inductance-capacitance filters (LC)
  power supplies 676
Induction
  audio transformers 275
Induction loop 1535
Induction loop system
  Assistive listening systems 1550
Inductive coupling
  grounding and interfacing 1188
Inductive filters
  power supplies 675
Inductive loads
  relays 379
Inductive loop systems
  message repeaters and evacuation systems 1520
Inductive networks
  filters and equalizers 788
Inductive reactance 248, 265–267, 270, 864, 880, 891, 1183, 1548
  assistive listening systems 1548
  audio transformers 275–276, 285
  consoles 841, 876, 879, 888
  fundamentals of audio and acoustics 33
  of a capacitor 258
  power supplies 676
Inductor
  amplifier 713
  audio transformers 275
  grounding and interfacing 1183
Inductors 263
  air core 263
  air gap 266
  axial 263
  bobbin core 263
  ceramic core 263
  coil inductance 266
  color code 263
  core losses 267
  epoxy coated 264
  ferrite beads 268
  ferrite core 264
  flux density 266
  impedance 265
  inductance 263
  inductive reactance 265–267, 270
  insertion loss 266
  laminated core 264
  magnetic field 263, 266, 268
  molded 264
  MPP core 264
  multilayer 264
  mutual inductance 265
  phenolic core 264
  powdered iron core 264
  Q 267
  radial inductor 265
  right-hand rule 266, 268

shielded 265, 268
skin effect 268
slug core 265
tape wound core 265
time constant 267
toroidal 265
types of inductors 263
winding 266
Industrial intercom systems 1566
In-ear monitoring
personal monitors 1416, 1420
Infinite baffle
loudspeakers 633
Infinite impulse response (IIR) linear filter
DSP technology 1167
Information/broadcast services applications
message repeaters and evacuation systems 1518
Infrared control transmitter
ssistive listening systems 1553
Infrared headset 1538
Infrared link
assistive listening systems 1553
Infrared power
interpretation and tour group systems 1538
Infrared signal
interpretation and tour group systems 1538
Infrared slave emitters
assistive listening systems 1553
Infrared systems
assistive listening systems 1552
interpretation and tour group systems 1536
message repeaters and evacuation systems 1520
Infrared systems channel allocation
assistive listening systems 1555
Infrared systems modulation technique
assistive listening systems 1555
Initial Time Gap
delay 808
Injection laser diode (ILD)
fiber optics 481
Inline microphone processors 573
Inner ear
psychoacoustics 43, 45
Input capacitance
tubes 313
Input capacitance compensation
integrated circuits 338
Input channel attenuation
preamplifiers and mixers 747
Input coupling
consoles 872
Input impedance
consoles 865
preamplifiers and mixers 737
Input resistance
transistor 328
Input saturation
consoles 842
Input transformer
audio transformers 288
Input transformers
audio transformers 289
Inputs
consoles 976
Inrush current
relays 375
Insertion loss
attenuators 769, 773
audio transformers 286, 290
fiber optics 471–472, 475, 481
inductors 266
relays 375, 385
Instability
consoles 838
Installation
audio transformers 306
Institute of Electrical and Electronic Engineers 12
Institute of Radio Engineers 12
Instruction cycles
consoles 959
Instrumentation amplifier
integrated circuits 358
Insulated-gate transistor
solid state devices 327
Insulation
wire and cable 403
Insulation characteristics
wire and cable 405
Insulation color codes
wire and cable 407
Insulation resistance
relays 375
Integrated circuit preamplifiers
integrated circuits 349
Integrated circuits 334
active balanced line driver 360
active balanced line receiver 360
active microphone preamplifier 350
AD524 350
AD636 347
audio dB meter 348
average level detection 345
average-responding level detection 345
balanced bridge 353
balanced line input circuits 356
balanced line interfaces 353, 355
balanced line outputs 354
balanced line receiver 356–357
bipolar junction transistor 343
CMMR 340
CMR 354
common-mode impedance 354
common-mode load impedance 354
common-mode rejection (CMR) 354, 357
common-mode rejection line receiver 358
differential amplifiers 339
differential signal 353
digital 360
dual balanced line receiver 355
Fairchild μA741 340
gain and bandwidth 338
hybrid 335
input capacitance compensation 338
instrumentation amplifier 358
integrated circuit preamplifiers 349

- inverting amplifiers 338
- monolithic 334
- MOSFET transistors 335
- National LM381 340
- noninverting amplifiers 338
- op-amps 344, 358
- operational transconductance amplifier 340
- operational voltage amplifiers 335
- output capacitance compensation 338
- peak level detection 345
- power supply compensation 338
- rms level detection 345
- single ended bootstrap 358
- SSM2018 343
- SSM2141 355
- state-variable filter 346
- sum-difference networks 356
- summing inverter amplifiers 340
- THAT 2180 343
- THAT1200-series 355
- THAT1510 351
- THAT2252 347
- transformer input microphone preamplifiers 349
- VCA 341
- voltage controlled amplifier 342

Integrated control of outboard gear
- consoles 953

Integrated detector preamplifier
- fiber optics 467

Integrated Optical Components (IOCs)
- fiber optics 481

Integrated sequencers
- MIDI 1118

Integration times
- psychoacoustics 54

Intelligibility
- %Alcons 1391–1392, 1398, 1400, 1404, 1408
- Alcons 157, 159
- comb filters 1390
- designing for speech intelligibility 1397
- loudspeaker clusters 654

Intelligibility criteria
- designing for speech intelligibility 1403

Intelligibility measurement
- designing for speech intelligibility 1403

Intelligibility of sentences
- acoustics for auditoriums and concert halls 157

Intelligibility of words
- acoustics for auditoriums and concert halls 157

Intelligibility optimization techniques
- designing for speech intelligibility 1402

Intelligibility prediction
- designing for speech intelligibility 1395, 1400

Intensity
- fiber optics 481

Interaural Cross-Correlation Coefficient (IACC)
- acoustics for auditoriums and concert halls 153

Interaural intensity difference (IID)
- psychoacoustics 57

Interaural level difference (ILD)
- psychoacoustics 57

Interaural phase difference
- psychoacoustics 50

Interaural phase difference (IPD)
- psychoacoustics 57

Interaural temporal difference
- psychoacoustics 58

Interaural time difference
- psychoacoustics 57

Intercommunication systems
- intercoms 1561

Intercoms
- 4-channel, two wire party line system 1570
- absence registration 1566
- all call 1566
- apartment intercom systems 1563
- beltpacks 1575
- broadcast 1569
- broadcast matrix systems 1571
- broadcast party line systems 1570
- call back 1566
- call forwarding 1567
- call reply 1567
- call transfer 1567
- camp on 1567
- capture ratio 1575
- central answering service 1567
- commercial security systems 1564
- commercial systems 1566
- conference 1567
- conference line systems 1561
- crosspoint electronics 1572
- dedicated line 1561
- digital matrix systems 1561–1562, 1572
- digital potentiometers 1573
- display features 1567
- distributed line systems 1561
- DTMF 1563, 1575
- emergency call systems 1561
- emergency crisis communications 1566
- equipment locator 1568
- fiber optic systems 1572
- FM transmission 1575
- group callback 1567
- hands free 1567
- headsets 1575
- hospital facilities systems 1566
- industrial systems 1566
- intercommunication systems 1561
- interrupted feed-back 1573
- interrupted fold-back 1573
- interrupted return-feed 1573
- last number redial 1568
- matrix systems 1571
- mechanical switching matrix 1562
- microphone cutout 1568
- nurse call 1561
- nurse call systems 1568
- paging 1568
- party line system 1570
- party line systems 1561
- PBX lines 1563
- PBX station intercommunica-

tions 1568
pocket page access 1568
point-to-point systems 1561, 1570
priority 1568
privacy 1568
private line 1561
program distribution channels 1568
program interrupt (IFB) 1570, 1573
remote Mic-Kill 1575
residential intercom systems 1564
satellite links 1572
scan monitoring 1568
security audio monitoring 1564
security intercom systems 1563
telephone interface 1575
two channel party-line system 1570
two-wire party line matrix system 1571
voice over IP 1572
VOIP 1566
wireless 1569, 1574
zone paging 1564
Inter-driver spacing versus wavelength
loudspeaker clusters 659
Interface
MIDI 1113
Interfaces
Gain structure 1223
grounding and interfacing 1185
Interference patterns
loudspeaker clusters 648
Interference suppression 261
Interference tube microphone 532
Interlaced scanning
display technologies 1581
Intermediate Cross-Connect (IXC)
wire and cable 412
Intermodal distortion
fiber optics 481
Internal capacitance
transistors 331
tubes 313
Internal output impedance
power supplies 672
Internal winding resistance
audio transformers 283
International power
wire and cable 445
Inter-outlet ground voltage noise
grounding and interfacing 1194
Interpolation
optical disc formats 1140
Interpolation filter
consoles 958
Interpretation and tour group systems
acoustical crosstalk 1534
Cat-5e cables 1531
central control unit 1531
Common Mode Reverse Audio Feed (CMRAF) 1534
FM interpretation systems 1534
FM receiver 1540
FM transmitter 1539
hard wired systems 1533
infrared power 1538
infrared signal 1538
infrared systems 1536
interpretation systems 1531
interpreter's booth 1531, 1533
ISO 4043 1981 1531
language interpreter's unit 1532
light emitting diode (LED) 1537
multichannel FM induction loop 1535
Soundfield FM applications 1541
tour group FM system 1538
Interpretation systems
interpretation and tour group systems 1531
Interpreter's booth
interpretation and tour group systems 1531, 1533
Interrupted feed-back
intercoms 1573
Interrupted fold-back
intercoms 1573
Interrupted return-feed 1573
Inter-stage coupling transformers
audio transformers 293
In-the-Ear recording (ITE) microphone 554–555
Intrinsic jitter
digital audio interfacing and networks 1471
Inverse Fast Fourier Transform
test and measurement 1615
Inverse square law
sound system design 1240
Inverse-frequency noise
consoles 860
Inverse-square law
fundamentals of audio and acoustics 37
Inverters
power supplies 690
Inverting amplifiers
consoles 869
integrated circuits 338
Inverting operational amplifier 710
Inverting voltage amplifier 711
IR LED transmitting diode
assistive listening systems 1554
Irradiance
fiber optics 481
Isaac Newton 5–6
ISDN circuits 1572
ISO 4043 1981
interpretation and tour group systems 1531
ISO 9660
optical disc formats 1146
ISO numbers
units of measurement 1662
ISO/IEC 10149 standard
optical disc formats 1141
ISO/IEC 13346
optical disc formats 1146
Isochronous connections
digital audio interfacing and networks 1462
Isochronous data packet
digital audio interfacing and networks 1505
Isolated grounding
grounding and interfacing 1214
Isolating earphone

- personal monitors 1433
- Isolation
  - audio transformers 290
  - relays 375, 385
- Isolation mount systems
  - acoustical noise control 80
- Isolation systems
  - acoustical noise control 74, 89
- ITG
  - delay 808

J

- J. Willard Gibbs 6
- Jackets
  - wire and cable 405
- Jackfields
  - consoles 952
- Jacks
  - MIDI 1111
- James Clerk Maxwell 5–6
- J-bend relays 386
- Jensen-Hardy Twin-Servo® 990 microphone preamplifier
  - preamplifiers and mixers 741
- Jim Lansing 11
- Jim Noble 12
- Jitter
  - digital audio interfacing and networks 1471
- John Blackburn 11
- John Hillard 12
- John Hilliard 9, 11
- John Volkmann 13
- Joseph Henry 6
- Just-noticeable difference (JND)
  - psychoacoustics 54
- Just-tuning
  - psychoacoustics 55

K

- K factors for calculating attenuator loss 768
- K value
  - attenuators 768
- Karl Kupfmuller 10
- Keying
  - consoles 906
- Kinetic Energy (T) 1649
  - units of measurement 1649
- Kirchhoff's current law (KCL)
  - units of measurement 1654
- Kirchhoff's laws
  - units of measurement 1653
- Kirchhoff's voltage law (KVL)
  - units of measurement 1653
- Klipschorn
  - loudspeakers 615

L

- L pads 771
- L. F. von Helmholtz 14
- Ladder attenuators 776
- Ladder DACs
  - consoles 957
- Ladder networks 776
  - filters and equalizers 791
- Ladder potentiometers 776
- LAN cabling
  - wire and cable 416
- Language interpreters unit
  - interpretation and tour group systems 1532
- LaPlace transfer function
  - loudspeakers 628
- Laplace transform 705
- LARES System
  - acoustics for auditoriums and concert halls 194
- Large area multi-channel systems
  - assistive listening systems 1553
- Large area systems (LAS)
  - assistive listening systems 1553
- Laser
  - fiber optics 465
- Laser diode
  - fiber optics 465, 481
- Laser turntable 1034
  - analog discs 1034
- Last number redial
  - intercoms 1568
- Latching relay 376, 380
- Late reflections
  - designing for speech intelligibility 1394
- Latency
  - consoles 991
  - personal monitors 1431
- Lateral efficiency LE
  - acoustics for auditoriums and concert halls 161
- Launch angle
  - fiber optics 481
- Launching fiber
  - fiber optics 481
- Lavalier microphone 527
- Law of induction
  - audio transformers 275
  - grounding and interfacing 1183
- Layering
  - audio transformers 281
- LC fiber optic connector 472
  - fiber optics 474
- LCD
  - display technologies 1584
- LCD projection TV
  - display technologies 1585
- LCoS
  - display technologies 1587
- LCoS projection system
  - display technologies 1587
- Lead-acid batteries 694
- Lead-dioxide batteries 695
- Leadership in Energy and Environmental Design (LEED)
  - acoustical treatment for indoor areas 122
- Leadless relays 386
- Leaf tweeter
  - loudspeakers 601
- Leakage current noise
  - grounding and interfacing 1194
- Leakage current of a capacitor 258
- Leakage inductance
  - audio transformers 281
- Least significant bit (LSB)
  - digital audio interfacing and networks 1464
- LeClanche (Zinc-Carbon) batteries 697
- Lectern microphone 589
- LED
  - display technologies 1588
  - fiber optics 463
  - lifetime
    - fiber optics 467
  - solid state devices 325
  - speed
    - fiber optics 467
- LEDE
  - small room acoustics 141
- LEDR (Listening Environment Di-

agnostic Recording
  psychoacoustics 59
Left-center-right cluster loudspeaker system
  sound system design 1261
Length (l)
  units of measurement 1648
Leo Beranek 15
Leq
  what's the ear for? 1640
Level
  fundamentals of audio and acoustics 26
Level control
  virtual systems 1439
Level difference
  fundamentals of audio and acoustics 26
Level matching
  virtual systems 1442
Level specifications
  consoles 834
LFE
  surround sound 1595
Light
  fiber optics 452, 481
Light activated silicon controlled rectifier
  solid state devices 325
Light emitting diode (LED
  assistive listening systems 1553
Light Emitting Diode (LED)
  display technologies 1588
Light emitting diode (LED)
  fiber optics 481
  interpretation and tour group systems 1537
Light prism
  fiber optics 453
Light propagation
  fiber optics 456
Light rays
  fiber optics 454
Light source specifications
  fiber optics 476
Light spectrum
  assistive listening systems 1555
Light-dependent attenuator (LDA) 778
Lightwaves
  fiber optics 481
Limiter
  consoles 874
Limiters
  personal monitors 1431
  sound system design 1279, 1286
Limiting
  consoles 895
Limiting distance
  units of measurement 1658
Limiting for effect
  consoles 904
Lincoln Walsh 13
Line array
  loudspeakers 620
Line array loudspeaker systems
  sound system design 1257
Line arrays
  Computer aided sound system design 1352
  fundamentals of audio and acoustics 38
  loudspeaker clusters 653, 655
Line driver
  digital audio interfacing and networks 1471, 1473, 1486
  grounding and interfacing 1187
Line impedance
  VU meters 1001
Line input transformer
  audio transformers 291
Line level inputs
  consoles 867, 873
Line Level Shifter
  audio transformers 303
Line receivers
  digital audio interfacing and networks 1471, 1473
Line source
  fundamentals of audio and acoustics 37
Linear amplifier 703
Linear phase
  filters and equalizers 798
Linear speed
  analog discs 1023
Linear time invariant
  filters and equalizers 786
Linear time invariant systems
  DSP technology 1164
Line-level output transformer
  audio transformers 292
Lines of force
  audio transformers 275
Linkwitz-Reilly filters
  loudspeakers 623
Linkwitz-Riley filter
  filters and equalizers 792
Liquid Crystal Display (LCD)
  display technologies 1584
Liquid Crystal on Silicon (LCoS)
  display technologies 1587
Listening rooms
  small room acoustics 140
Literal dynamics
  consoles 972
Load impedance
  Audio transformers 297
Load line
  transistors 333
Load lines
  transistors 334
Load protection mechanism 719
Loading coil
  wire and cable 431
Loading loss
  preamplifiers and mixers 735
Localization cues
  psychoacoustics 57
Location of HVAC equipment
  acoustical noise control 90
Log-antilog cell
  consoles 938
Logarithmic amplifier 703
Logarithmic scale
  fundamentals of audio and acoustics 23
  psychoacoustics 52, 55
Logatoms
  acoustics for auditoriums and concert halls 157
Logic control
  consoles 923
Logic sense
  consoles 928
Longitudinal digital tape recorders 1079

Look up table (LUT)
  digital audio interfacing and networks 1493
Loop Current 1548
Loop filters
  consoles 886
Loop impedance
  assistive listening systems 1548
Loop networking
  consoles 992
Loose tube fiber jackets 457
Lord Kelvin 5
Lord Rayleigh 14, 1339
  small room acoustics 127
Loss
  attenuators 769
Lossless encoders
  psychoacoustics 48
Lossy coaxial cable antenna
  assistive listening systems 1552
Lossy ladder network
  microphones 500
Lou Burroughs 12
Loudness
  designing for speech intelligibility 1390
  fundamentals of audio and acoustics 26
  psychoacoustics 51–52
Loudness control 776
Loudness meter
  VI meters 1004–1005
Loudness meters
  VI meters 1004
Loudness versus bandwidth
  psychoacoustics 53
Loudspeaker
  cable 435
  coincident acoustical centers 650
  control-room monitor loudspeaker 826
  low frequency arrays 652
  TRAP horns 650
  True Array Principle (TRAP) 651
  vertical woofer arrays 653
Loudspeaker array problems
  loudspeaker clusters 647
Loudspeaker arrays
  computer aided sound system design 1350
Loudspeaker balloon measurements
  computer aided sound system design 1369
Loudspeaker cable
  wire and cable 435
Loudspeaker clusters
  %Alcons 654
  ALS-1 interference predictions 649, 651
  array height versus wavelength 659
  Articulation Index (AI) 654
  beam steering 656, 662
  beneficial destructive interference arrays 653
  beneficial interference 652
  coincident acoustical centers 650
  column arrays 656
  column loudspeakers 661
  conventional array shortcomings 647
  coverage angle 648
  coverage boundaries 649
  destructive acoustic interference 652
  digitally steered column arrays 656
  doublet source directivity 658
  DSP-driven line arrays 662
  DSP-driven vertical arrays 657
  high frequency beaming 656
  horizontal directivity 663
  horizontal woofer arrays 652
  intelligibility 654
  inter-driver spacing versus wavelength 659
  interference patterns 648
  line arrays 653, 655
  loudspeaker array problems 647
  low frequency arrays 652
  multichannel DSP 661
  point source interactions 658
  progressive delay drivers 663
  reference point array 653
  reverberation 655
  Speech Transmission Index (STI) 654
  steerable loudspeaker column arrays 653
  TRAP horns 650
  tri-polar array 653
  true array principle (TRAP) 651
  tuned dipolar array 653
  variable Q 662
  vertical woofer arrays 653
Loudspeaker components
  loudspeakers 597
  sound system design 1255
Loudspeaker connectors
  sound system design 1295
Loudspeaker failure modes
  sound system design 1276
Loudspeaker matching transformers
  audio transformers 300
Loudspeaker measurements
  computer aided sound system design 1357
Loudspeaker polar data
  computer aided sound system design 1365
Loudspeaker power capacity specifications
  sound system design 1277
Loudspeaker power ratings
  gain structure 1231
Loudspeaker processing DSP
  sound system design 1288
Loudspeaker protection devices
  sound system design 1278
Loudspeaker Q
  sound system design 1246
Loudspeaker response
  loudspeakers 623
Loudspeaker sensitivity
  units of measurement 1661
Loudspeaker systems
  70.7 volt loudspeaker system 435
  computer aided sound system design 1343
  sound system design 1255, 1260
Loudspeakers
  acoustic boundaries 625
  acoustic lens 614

acoustic power 633
aiming
  computer aided sound system design 1374
air gap length 608
air mass 635
amplitude response 628, 631
angular directivity ratio 1347
asymmetric-directivity horns 613
baffle 626
bass reflex boxes 1350
Bessel array 208
Bessel function 634
Biflex 598
bimorph bender 604
binaural reproduction 232
biradial 613
cabinet diffraction 626
closed-box system 636
coaxial 620
coil/gap geometries 610
common loudspeakerformat 629
communication 597
components 597
compression drivers 600
cone radiator 610
cone type 603
constant directivity horns 612–613
constant voltage crossovers 624
crossover frequency 623
crossover network 622
crossover systems 809
cutoff frequency 638
damping 599
desktop loudspeakers 918
diaphragm types 598
digitally controlled line arrays
  computer aided sound system design 1352
Dirac delta function 640
direct radiator 610, 626
direct radiator enclosure design 635
direct sound radiation 633
directivity data 632
distortion 629
dome radiator 610
dome tweeter 626
DSP technology 621
dual channel FFT measurement 640
efficiency 609, 627
  computer aided sound system design 1347
electrodynamic transducers 598
electrostatic 603–604
electrostatic transducers 602
enclosure spring constant 635
explosion proof drivers 602
ferrofluids 608
filter slopes 623
flux density 636
folded horns 614
Fourier transform 640
Frazier 620
freestanding radiator 635
frequency resolution 639
frequency response 628
fundamentals of audio and acoustics 36
heat conduction 608
heat transfer designs 609
heat transfer mechanisms 609
highpass filter 622, 635
horn 1350
horn directivity 638
horns 611
impedance 629
impulse response 624, 628
infinite baffle 633
Klipschorn 615
LaPlace transfer function 628
leaf tweeter 601
line array 620
line arrays
  computer aided sound system design 1352
  fundamentals of audio and acoustics 38
line source 37
Linkwitz-Reilly filters 623
lowpass filter 622
maximum length sequence (MLS) 640
mechanical failure 609
motor 598
motor design 605
multicell horns 612
multichannel reproduction
  acoustical modeling and auralization 233
narrow beam
  message repeaters and evacuation systems 1524
near-field 918
outdoor sound
  arraying devices 206
  Bessel array 208
  coverage 203
  distributed loudspeaker systems 210
  medium and long throw devices 205
  parabolic loudspeaker 209
overhung coil 607
panel radiators 611
performance data 631
performance definitions 631
pest damage 1280
phase plug 601
phase reversal boxes 1350
piezoelectric 604
piezoelectricity 604
pink noise 639
piston directivity 634
piston resistance 633
polar data
  computer aided sound system design 1365
Proportional Point Source 614
protecting against weather 1280
radial horns 612
radiator types 609
realtime analyzers 639
residential music systems 618
response 623
ribbon 606
ring radiator 611
sealed-box system 635
sensitivity 627
servo-drive 606
signal delay 1307
sound pressure level 627
sound production 597
sound reproduction 597
spectrum analysis 638

  spider 598–599
  surround 598
  suspension methods 599
  swept sine measurement 641
  thermal failure 609
  thermal limit 607
  thermal resistance 608–609
  Thiele-Small parameters 637
  Time Delay Spectrometry 641
  time windowed measurements 639
  transducer types 598
  transfer function 624, 629
  two-way crossovers 624
  underhung coil 607
  underwater 605
  vented boxes 636
  virtual systems 1453
  viscoelastic 599
  voice coil 598
  voice warning horns 613
Loudspeakers for distributed systems
  sound system design 1270
Low frequency arrays
  loudspeaker clusters 652
Low frequency response 1057
  acoustical noise control 69
Low load current protection
  relays 390
Low pass filter 714
  filters and equalizers 785, 789
Low Power Radio Services (LPRS)
  Assistive listening systems 1550
Low RT60 rooms
  sound system design 1254
Low thermal Emf relay 376
Low-frequency absorption
  computer aided sound system design 1354
Low-frequency control
  consoles 894
Low-level logic switching
  relays 382
Low-loss LC filter
  preamplifiers and mixers 736
Lowpass filter
  DSP technology 1166
  loudspeakers 622
Lowpass filtering
  DSP technology 1162
Low-pass filters
  sound system design 1288
LP records 1015
LPCM
  optical disc formats 1148
Lumen
  display technologies 1579
Luminance
  display technologies 1582
Luminous flux
  units of measurement 1651
Luminous flux density
  units of measurement 1651
Luminous Intensity (IL)
  Units of measurement 1649

M

M/S stereo
  microphones 544
Macrobending
  fiber optics 481
Magnetic circuits
  audio transformers 280
Magnetic coupling
  audio transformers 288
Magnetic Disks 1083
Magnetic field 1027, 1546
  amplifier 703–704
  analog discs 1027–1028
  audio transformers 275
  cartridges 1027
  grounding and interfacing 1183–1185, 1188
  in inductors 263, 268
  inductors 266
  magnetic field 1044
  magnetic fieldding 1053, 1060, 1064, 1066–1067, 1072, 1094
  magnetic recording 1053, 1069, 1083, 1085
  power supplies 676
  relays 380–381
magnetic field
  magnetic recording 1085
Magnetic field strength
  assistive listening systems 1546, 1548
Magnetic flux 277
  units of measurement 1651
Magnetic Flux (f) 1651
Magnetic flux density
  audio transformers 284
  units of measurement 1651
Magnetic Flux Density (ß) 1651
Magnetic head troubleshooting 1094
Magnetic heads 1053
Magnetic hysteresis
  audio transformers 279
Magnetic induction loop system
  assistive listening systems 1546
Magnetic induction loops
  assistive listening systems 1546
Magnetic interaction
  relays 381
Magnetic operating point
  audio transformers 279
Magnetic Performance Curves 1067
Magnetic recording 1041, 1053, 1057–1060, 1063–1064, 1066–1067, 1072, 1091, 1093–1094
  3M Isoloop drive. 1049
  anhysteretic recording 1062
  automated alignment 1097
  base film materials 1065
  bias circuits 1078
  biased recording 1062
  bulk eraser 1063
  capstan speed control 1045
  capstan-derived tensioning 1048
  constant current record head drivers 1076
  contour effect 1057
  DASH recorders 1079
  degausser 1063
  digital recording 1083
  digital signal processor 1086
  dual winding record head 1063
  eddy currents 1058–1059
  equalization boosts 1056
  erase circuits 1078
  floppy disks 1083
  flux 1051, 1053

flux loop 1094
flying head 1085
gap jumping 1064
giant magnetoresistive heads 1082
hard disk drives 1082, 1084
hard disk electronics 1086
head relapping 1095
helical-scan digital tape transports 1080
high level formatting 1087
inductive rise 1058
laminations 1059, 1062, 1088
longitudinal digital tape recorders 1079
long-term tape storage 1087
low level formatting 1087
magnetic core 1042
magnetic field 1053, 1069, 1083, 1085
magnetic head troubleshooting 1094
magnetic heads 1053
magnetic recording 1055–1058
magnetic recording currents 1061
magnetic recordingad degaussing 1064
magnetic recordingarkhausen noise 1060
magnetic recordingazimuth 1051
magnetic recordingeed 1089
magnetic recordingion 1063–1064
magnetic recordingk 1058
magnetic recordingn 1056
magnetic recordingonse 1057
magnetic recordingp length loss 1053
magnetic recordings 1074
magnetic recordingstems 1078
magnetic tape 1041
magnetic tape specifications 1068
maintenance 1088
pinch roller 1046
pinch roller wear patterns 1090
plastic substrate 1071
ProDigi machines 1079
R-DAT 1080
record circuits 1076
rigid disks 1084
rotary digital audio tape 1080
scrape flutter 1052
servos 1086
signal alignment procedure 1095
sliding friction 1053
spin valve heads 1082
spooling-motor-derived tensioning 1049
squealing tape 1074
static friction 1053
step-up transformer 1063
sync operation 1079
tape adhesion 1071
tape binder breakdown 1073
tape brittleness 1071
tape guides 1052
tape guiding 1050
tape height 1051
tape lubricant oozing 1073
tape metering 1043
tape peeling oxide 1071
tape record-playback equipment equalization standards 1056
tape splice failures 1071
tape tranports 1043
THD 1059
timebase corrector 1080
timecode 1087
track width 1060
variable speed oscillator 1044

Magnetic saturation 1056
  audio transformers 283
Magnetic shield
  audio transformers 283, 301
  relays 376
Magnetic shielding
  audio transformers 283
Magnetic tape 1041
Magnetic wire
  audio transformers 282
Magnetizing material
  audio transformers 278
Maintenance
  audio transformers 306
Make or break current
  relays 377
Malleus
  psychoacoustics 43
Managed switches
  digital audio interfacing and networks 1495
Manual/graphical cluster design tools
  sound system design 1264
Manufaacturer
  Optical Radiation Corp. 1595
Manufacturer
  3Com Corporation 470
  Ableton 1122
  AcousticX 129
  Acoustiguide Inc 1521
  Adamson Acoustics 599
  Agilent Technologies 469
  Aiphone Communication 1563
  Aiphone Communications 1563
  AKG Acoustics, Inc. 509, 545
  ALARMCO 1517–1519
  Alesis 1118
  Allen and Heath 1281, 1283
  Altec Lansing 749
  Altec Lansing Corp. 439, 598, 601, 612–613
  AMX 1441
  Analog Devices, Inc 351
  Analog Devices, Inc. 342
  Atlas Sound 581, 1271, 1330, 1332
  ATM Hardware 1313
  Audio-Technica U.S., Inc. 557
  Auditel Systems Limited 1532, 1535
  Auralex Acoustics, Inc. 106, 122
  Belden 401, 404, 414, 426, 436, 736
  Belden Wire 1217
  Biamp 1447
  Biamp Systems 748–749
  Bogen 1519
  Bogen Communications, Inc. 1561
  Bosch/Electro-Voice 1257, 1264, 1269, 1311, 1322,

1325
Brüel & Kjaer Instruments, Inc. 1309
Bruel and Kjaer 9
Brüel and Kjaer 70–71, 576, 1407, 1614
BSS 1290, 1441, 1446
Cirrus Logic, Inc. 1504
Clear-Com Communication Systems 1571, 1574
Clevite Corp. 503
Coil-Craft 1214
Community Professional 1260, 1262, 1328
Community Professional Loudspeakers 599, 602, 613
Cooper Notification 1526–1527
Cooper Sound Systems 740
Corning, Inc. 474
Coto Technology 381, 383, 385–386
Countryman Associates, Inc. 528–529
Crestron 1441, 1448
Crown 1441
Crown International 1289–1290
Crown International, Inc. 523, 550
Dan Dugan Sound Design 749, 755
DataKustik GmbH 231
dbx Inc 341
dbx Professional 1280, 1286
DIGIPRESS 1527
Dolby Laboratories 341, 1593
Dorrough Electonics 1003
Dorrough Electronics 36, 1006
DPA Microphones, Inc. 534, 564
EAW 652, 1256
EAW/SIA Software 1309
Electro-Voice, Inc. 581
ElectroVoice, Inc. 613
ELP Corporation 1034–1035
Espro 1522
Eventide 1440
Fender Musical Instruments 1296–1297, 1299
Fluke Networks 462
Frazier Loudspeakers 620, 632
Future Sonics 1423
Galaxy Audio 1608
Genelec 619
Georg Neumann GmbH 554
Gold Line 68, 70, 1308, 1608
Hewlett Packard 18
Holophone® 552
Holosonic Research Labs, Inc. 1525
IAC 85, 88
Innovative Electronic Designs 1447
JBL 600, 613, 615
JBL Professional 1246–1247, 1254, 1266, 1271, 1312
Jensen Transformers 358
Jensen Transformers, Inc. 304, 741–742, 1200, 1208
KEF Company 599
Kinetics Corp 80
Kinetics Corp. 81
Kirkegaard Associates 216, 218
Klipsch & Associates, Inc. 1256
Kodak 1595
L-Acoustic 652
Lectrosonics 748–749, 752, 1285
Lectrosonics, Inc. 574
Level Control Systems 1444
Lexicon 194
Libbey-Owens-Ford Co 85
Linear Technology Corporation 684–685
Listen Technologies Corporation 1531, 1533, 1539, 1551
Loud Technologies, Inc. 1124
Lowell 1271, 1292
Lubell Laboratories, Inc. 605
Lucent Technologies 475
Mackenzie Laboritories, Inc. 1519
Magnecraft Electric Co. 381–382
Marshall Electronics 575
Marshall Electronics, Inc. 557
Mason Industries 74, 79, 81–82
M-Audio 1118
Maxim Integrated Products 685, 687–688
Maxxon 80
Meyer Sound 653
Meyer Sound Laboratories, Inc. 209
Mitsubishi Electric 599–600
Motorola, Inc. 604
Neumann USA 533
Neutrik AG. 473
Neutrik USA, Inc. 1295
Novation Digital Music Systems, Ltd 1117
Onlife Research, Inc. 1026
Owens Corning 103
Owens-Corning Fiberglas Corp 92
Panasonic Industrial Corp. 601
Panduit Int'l Corp. 417
Peak Audio 1488, 1504
Peavey 1289, 1446–1447
Peavey Electronic Corp. 600
Pinta Acoustic, Inc. 106
Polycom, Inc. 754–756
Propellerheads 1123
Protech Audio 749
Proudfoot 111
QSC 1289–1290, 1447, 1449, 1454
Quest Technologies 70
Ram Technologies LLC 691
Rane 1287, 1291, 1442
Rauland Borg 1568
Renkus Heinz 651, 1258, 1265
Ring Communications, Inc. 1565, 1567
RPG Diffuser Systems Inc. 1354
RPG, Inc. 110, 113, 119
RTW GmbH & Co.KG, Cologne 1004–1005, 1008
Rycote Microphone Windshields LTD 577, 580
Sabine, Inc. 574

Schall-Technik 511
Schoeps GmbH 541, 551
sE Electronics 578
Sennheiser 1416, 1419
Sennheiser El ectronic Corporation 1552
Sennheiser Ele ctronic Corporation 1557
Sennheiser Electronic Corporation 501, 509, 514, 532
Sennheiser Electronics 1520, 1523, 1538
Sensaphonics 1423
Shure 1419
Shure Incorporated 499–500, 529, 531, 540, 579, 735, 740, 750, 761, 1416, 1419
Shure Incorporated. 751
Simpson Electric 998
Sound Control Technologies, Inc. 761–762
SoundPLAN LLC 231
Stanton Magnetics, Inc. 1034
Symetrix 1325
Tectum, Inc. 104
Tektronix® 462
Telex Communications, Inc. 1575
Telex Electro-Voice 497
Texas Instruments 351, 355
THAT Corporation 343, 353, 355
TOA 1441, 1445
Tour-Mate 1522
Ultimate Ears 1423
Vishay Dale 268
Vishay Roederstein 261
VPI Industries 1037
VPI Industries, Inc. 1021, 1037
Wheatstone Corporation 950, 994
Whelen Engineering 614
Williams Sound 1523
Yamaha 1282, 1447
Yamaha Corporation 194
Yamaha Corporation of America 1114
Yamaha International Corp. 598, 601, 611, 1294, 1303, 1305
Mapping
  computer aided sound system design 1378
Mapping dynamics
  consoles 971
Marcus Vitruvius 1339
Marginal relay 380
Masker
  psychoacoustics 48
Masking
  psychoacoustics 48
Mass (M)
  units of measurement 1648
Master functions
  consoles 932
Matched interface
  gain structure 1223
Matrix crosstalk
  consoles 854
Matrix mixing
  sound system design 1285
Matrix systems
  intercoms 1571
Max program level
  units of measurement 1661
Maximum continuous current rating
  wire and cable 402
Maximum length sequence (MLS)
  loudspeakers 640
  test and measurement 1616
Maximum peak current
  power supplies 669
Maximum peak inverse voltage
  power supplies 669
Maximum power output sensitivity
  microphones 516
Maximum working voltage
  capacitors 255
Mean free path
  small room acoustics 135
Mean free path (MFP)
  units of measurement 1660
Mean reverberation time
  acoustics for auditoriums and concert halls 149
Measurement parameters for typical balloon resolutions
  computer aided sound system design 1358
Measurement system
  test and measurement 1613
Measurements 1647
  audio transformers 302
Mechanical failure
  loudspeakers 609
Mechanical recording
  consoles 826
Mechanical switching matrix
  intercoms 1562
Media assess control (MAC)
  digital audio interfacing and networks 1488
Media converter
  digital audio interfacing and networks 1495
MediaMatrix
  virtual systems 1447
Medium area systems (MAS)
  assistive listening systems 1553
Medium considerations
  acoustical modeling and auralization 217
Medium density fiber panels
  acoustical treatment for indoor areas 103
Medium-frequency horns
  computer aided sound system design 1350
Mel
  psychoacoustics 55
Melamine foam panels
  acoustical treatment for indoor areas 106
Melodic intervals
  psychoacoustics 55
Melody
  psychoacoustics 55–56
Membrane absorbers
  acoustical treatment for indoor areas 109
MEMS switches
  relays 387
Mercury dry cell batteries 699
Mercury-wetted contact relays 381, 384
Mercury-wetted reed relays 384
Meridian Lossless Packing (MLP)

  Optical disc formats 1151
Mesa transistors
  solid state devices 325
Meshing
  digital audio interfacing and networks 1497
Message repeaters 1516–1517
Message repeaters and evacuation systems
  audio archival 1527
  digital audio storage 1515
  digital on-hold players 1519
  digital signal processors 1515
  directivity 1524
  DRAM 1528
  EEPROM 1528
  EPROM 1528
  factory floor applications 1517
  hospital applications 1516
  inductive loop systems 1520
  information/broadcast services applications 1518
  infrared systems 1520
  message-on-hold applications 1527
  museum applications 1520
  narrow beam loudspeaker system 1524
  retail store applications 1517
  RF systems 1520
  ROM 1528
  school applications 1518
  STI 1526
  transportation applications 1518
  UFC 1526
  ultrasound 1524
  voice evacuation 1526
  voice evacuation/mass notification systems 1526
Message repeaters operation 1527
Message-on-hold applications
  message repeaters and evacuation systems 1527
Metal oxide-semiconductor capacitor
  delay 810
Metal-oxide varistors (MOV)
  relays 389
Metal-oxide-semiconductor junction FET capacitor
  delay 810
Metering
  consoles 980
Meters
  consoles 919, 933
MFLOPS
  DSP technology 1176
MFP
  small room acoustics 135
MGM 14
Mic Mate USB adapter
  microphones 575
Mic'ing techniques 584
Michael Faraday 5–6
Michael I. Pupin 7
Michael Pupin 8
Microalloy diffused transistor
  solid state devices 325
Microbend loss
  fiber optics 459
Microbending
  fiber optics 481
Microlensed LED
  fiber optics 463
Micron
  fiber optics 481
Microperforated materials
  acoustical treatment for indoor areas 112
Microphone
  psychoacoustics 45
microphone
  condenser microphone 295
Microphone accessories 573
Microphone as a signal source
  preamplifiers and mixers 735
Microphone booms 581
Microphone calibrator
  test and measurement 1610
Microphone floor stands 581
Microphone identification
  digital audio interfacing and networks 1478
Microphone impedance conversion
  audio transformers 295
Microphone input transformer
  audio transformers 290
Microphone mixing
  preamplifiers and mixers 759
Microphone output transformer
  audio transformers 294
Microphone placement
  test and measurement 1619
Microphone practices 519
Microphone preamplifier fundamentals
  preamplifiers and mixers 735
Microphone sensitivity 518
Microphone snake cables
  sound system design 1295
Microphone table stands 581
Microphone techniques 582
Microphone transformers
  consoles 861
Microphones

    RF frequencies 564
  above the source 520
  above versus underneath mic'ing 585
  accessories 573
  aluminum diaphragms 504
  ambient 1431
  antenna cables 569
  antenna combination diversity 562
  antenna placement 562
  antenna splitters 570
  antennas 568
  artificial head systems 553
  attenuators 581
  automatic microphone systems 540
  balanced 521
  base station 530
  behind objects 520
  bidirectional 493, 495
  bidirectional pickup pattern 495
  binaural recording 553
  booms 581
  capacitor 507
  capacitor radio-frequency 512
  capacitor radio-frequency, frequency-modulated microphone 511
  capture ratio and muting 560
  carbon 502
  ceramic 503
  characteristics 583

coincident 542
comb filters 521, 523
companding 563, 567
computer aided sound system design 1355
condenser 507, 533, 583
condenser microphone 582
controlled-reluctance 530
criteria for selecting a wireless microphone 558
crystal 503
cutout 1568
differential noise-canceling 530
diffuse-field equalization 555
digital wireless transmission 572
direct versus reverberant field 520
directional pattern 584
distance from large surfaces 519
diversity antenna 571
Dual Pro Pop pop screen 578
dynamic 503
EIA output sensitivity 517
electret 515
equalizers 581
first-order gradient 493
floor stands 581
frequency response 498
frequency response of a cardioid microphone 497
grounding 520
handheld entertainer 530
head worn 528
headphone diffuse-field equalizers 554
Holophone surround sound system 551
human external-ear transfer function 554
hyper-cardioid 495
impedance 522
inline processors 573
interference tube 532
intermodulation (IM) 565
In-the-Ear recording (ITE) 554–555
lavalier 527
lectern 589
lossy ladder network 500
M/S stereo 544
maximum power output sensitivity 516
Mic Mate USB adapter 575
mic'ing bass drums 585
mic'ing cymbals 585
mic'ing drums 584
mic'ing electric instruments 592
mic'ing group vocals 589
mic'ing horns 591
mic'ing musicians 584
mic'ing percussion 592
mic'ing piano 586
mic'ing singers 588
mic'ing strings 590
mic'ing techniques 584
mic'ing Tom-Toms 585
mic'ing vocals 588
mic'ing woodwinds 591
microphone sensitivities 518
microphone techniques 582
microphone-to-source distance 519
moving-coil 503, 506
multiple-entry cardioid 501
mylar diaphragms 504
noise reduction 514
noise sources 513
omnidirectional 493–494, 505
omnidirectional diaphragm 504
omni-directional pickup pattern 494
open-circuit voltage sensitivity 516
operation of zoom microphones 539
ORTF technique 544
parabolic 537
performance characteristics 493
phantom power for capacitor microphones 510
Phase Coherent Cardioid (PCC) 524
pickup patterns 493
plug-on transmitter 574
PolarFlex microphone system 540
polarity 521
pop filters 576
pop protection 578
practices 519
pressure 493, 503
pressure zone (PZM) 522
pressure-gradient condenser 531
proximity effects 497
PZM directivity 524
PZM direct-to-reverberant ratio effect 525
PZM frequency-response anomalies caused by boundaries 526
PZM frequency-response effects 525
PZM sensitivity effects 524
range of a wireless transmission 566
reflection filter 578
RF power output 560
RF signal strength 568
rifle 537
Sabine Phantom Mic Rider 573
SASS 550
sensitivity 516
shock mounts 578
shotgun 532
single-entry cardioid 499
spurious emissions 567
stands and booms 581
step-up transformer 503
stereo 542
stereo boom 548
stereo mic'ing techniques 582
super-cardioid 495
surround sound system 550
suspension compliance 579
switching diversity antennas 561
symmetrical push-pull transducer 512
system planning for multi-channel wireless systems 571
table stands 581
thermal noise 518
three-entry cardioid 500
transmitter antenna 567
true diversity antennas 561

  two-way cardioid 502
  types of cardioid microphones 499
  types of transducers 502
  unbalanced 521
  unidirectional 493, 495, 498, 506
  USB 556
  velocity 493
  virtual systems 1452
  voltage divider capacitor 508
  waterproof wireless systems 563
  wind noise reduction figures 576
  windscreens 576
  wireless audio processing 567
  wireless communication systems 557
  wireless distortion 560
  wireless dynamic range 560
  wireless frequency bands of operation 558
  wireless frequency response 560
  wireless microphone receiver sensitivity 560
  wireless monitoring systems 571
  wireless multi-channel applications 565
  wireless receiver 568
  wireless receiving antenna systems 561
  wireless transmitter considerations 566
  XY stereo 543
  zoom 538
  zoom-microphone video camera linkage 540
Microphone-to-source distance 519
Microphonic
  audio transformers 306
Microprocessor
  DSP technology 1172
Microsoft HDi Interactive Format
  optical disc formats 1153
Middle ear
  personal monitors 1432
  psychoacoustics 43, 45
Middle ear infection
  psychoacoustics 45
MIDI
  algorithmic composition programs 1123
  cable 1111
  channel voice messages 1104
  cueing messages 1129
  daisy chain 1112
  drum machine 1116
  editing 1120
  full messages 1129
  guitars 1118
  integrated sequencers 1118
  Interface 1113
  jacks 1111
  MIDI channels 1102
  MIDI messages 1101
  MIDI modes 1103
  Mode 1–Omni On/Poly 1104
  Mode 2–Omni On/Mono 1104
  Mode 3–Omni Off/Poly 1104
  Mode 4–Omni Off/Mono 1104
  most significant bit (MSB) 1102
  MTC conversion 1129
  multiport network 1113
  Musical Instrument Digital Interface 1101
  music-printing programs 1124
  patch editors 1123
  percussion controllers 1117
  phantom power 1111
  playback 1120
  Poly/Mono mode 1104
  quarter-frame messages 1129
  real-time messages 1127
  reception modes 1104
  recording 1119
  sampler 1115
  sequencers 1118
  SMPTE 1128
  SMPTE-to-MTC converter 1128
  software sequencers 1119
  standard MIDI files 1125
  synchronization 1126
  synthesizer 1114
  system messages 1108
  time code 1128
  wireless MIDI 1111
MIDI channels
  MIDI 1102
MIDI messages
  MIDI 1101
MIDI modes
  MIDI 1103
MIDI sequencing
  consoles 983
Mineral fiber board
  acoustical treatment for indoor areas 104
Minimizing noise
  consoles 853
Minimum audible
  psychoacoustics 51
Minimum phase
  filters and equalizers 786
Minimum surface area
  heatsinks 372
Minimum-gain
  consoles 874
Minimum-loss pad 769
Mix code
  consoles 977
Mix groups, auxiliary groups and matrix mixing
  sound system design 1281
Mix stage
  consoles 976
Mixdown mode
  consoles 923
Mixed Mode CD
  optical disc formats 1142
Mixers
  attenuators 780
  sound system design 1280
Mixing
  consoles 910
Mixing consoles
  sound system design 1280
Mixing for personal monitors
  personal monitors 1427
Mix-minus mixer
  sound system design 1285
MLP
  optical disc formats 1151
MLS
  test and measurement

1616
Mnaufacturer
  AMP Incorporated 464
Mobility
  personal monitors 1418
Modal bandwidth
  small room acoustics 131
Modal dispersion
  fiber optics 481
Modal distribution
  small room acoustics 131
Modal noise
  fiber optics 481
Modal potency
  small room acoustics 131
Modal room resonances
  small room acoustics 128
Mode 1–Omni On/Poly
  MIDI 1104
Mode 2–Omni On/Mono
  MIDI 1104
Mode 3–Omni Off/Poly
  MIDI 1104
Mode 4–Omni Off/Mono
  MIDI 1104
Model
  acoustical modeling and auralization 215
Modeling
  acoustical modeling and auralization 215
  auralization process 232
  time and distance considerations 217
Model-size dummy head
  computer aided sound system design 1341
Modes
  fiber optics 481
Modes in nonrectangular rooms
  small room acoustics 133
Modified Chemical Vapor Deposition (MCVD
  fiber optics 481
Modified Rhyme Test (MRT)
  designing for speech intelligibility 1403
Modulation reduction factor m(F)
  acoustics for auditoriums and concert halls 155
Modulation transmission function (MTF)
  acoustics for auditoriums and concert halls 155
Molecular Substance (n)
  units of measurement 1649
Momentary actuation
  relays 392
Monitor mixes
  personal monitors 1422
Monitoring
  consoles 826, 917
Mono mixes
  personal monitors 1422
Monochromatic
  fiber optics 481
Monochromatic light
  fiber optics 465
Monolithic integrated circuits 334
MOSC
  delay 810
MOSFET transistors
  solid state devices 327
MOSFETs
  consoles 851
MOSJC
  delay 810
Most significant bit
  delay 813
Most significant bit (MSB)
  digital audio interfacing and networks 1464
  MIDI 1102
Motor
  loudspeakers 598
Motor design
  loudspeakers 605
Motor Loads
  relays 379
Motorized pots
  consoles 946
Moving coil cartridges
  analog discs 1027
Moving digital audio around
  consoles 985
Moving magnet cartridges
  analog discs 1028
Moving-coil phono step-up transformer
  audio transformers 301
Moving-coil phonograph pickups
  audio transformers 291
MP3 file
  DSP technology 1161
MP3 players
  DSP technology 1161
MPEG audio compression
  DSP technology 1163, 1166
MPEG-1
  Optical disc formats 1149
MPEG-2
  optical disc formats 1149
MPEG-2 AAC
  optical disc formats 1148
MPX
  personal monitors 1428
MS Windows dynamic link library (DLL)
  Computer aided sound system design 1360
MSB
  delay 813
  MIDI 1102
MTC
  MIDI Time code 1128
MT-RJ fiber optic connector
  fiber optics 473
Multiband processing
  consoles 902
Multicast bundles
  digital audio interfacing and networks 1506
Multicell horns
  loudspeakers 612
Multichannel audio digital interface
  digital audio interfacing and networks 1484
Multichannel DSP
  loudspeaker clusters 661
Multichannel FM induction loop
  interpretation and tour group systems 1535
Multichannel magnetic-stripe formats
  surround sound 1593
Multichannel optical format
  surroound sound 1593
Multichannel reproduction
  loudspeakers 233
Multichannel Wireless Microphone

and Monitoring Systems 564
Multi-Channel-Reverberation (MCR)
  acoustics for auditoriums and concert halls 192
Multicluster system
  sound system design 1270
Multiconductor wire
  wire and cable 407
Multi-filar winding
  audio transformers 281
Multifunction devices
  virtual systems 1444
Multi-function digital signal processing system
  sound system design 1288
Multi-genre theaters
  acoustics for auditoriums and concert halls 163
Multimedia Home Platform (DVB-MHP) standard
  Optical disc formats 1154
Multimode distortion
  fiber optics 482
Multimode fiber
  fiber optics 482
Multimode fiber optic cable
  digital audio interfacing and networks 1490, 1502
Multimode graded index fiber
  fiber optics 456
Multimode step index fiber
  fiber optics 456
Multipair cables
  wire and cable 408
Multiple and submultiple prefixes
  units of measurement 1647
Multiple effect feeds
  consoles 830
Multiple loudspeakers
  fundamentals of audio and acoustics 31
Multiple memory spaces
  consoles 961
Multiple open microphones
  preamplifiers and mixers 742
Multiple-entry cardioid microphone 501
Multiple-loop current flow
  assistive listening systems 1548
Multiplexers
  consoles 945
Multiplier/accumulator (MAC)
  consoles 959
Multiplying DACs 940
Multipoint grounding
  grounding and interfacing 1192
Multiport network
  MIDI 1113
Multi-purpose halls
  acoustics for auditoriums and concert halls 163
Multiread drives
  optical disc formats 1143
Multitrack console 922
Multitrack routing
  consoles 930
Multivariable dependent automatic microphone mixers 751
Mu-metal
  relays 382
Mumetal®
  audio transformers 283
Museum applications
  message repeaters and evacuation systems 1520
Musical Instrument Digital Interface
  MIDI 1101
Music-printing programs
  MIDI 1124
Mutual inductance
  inductors 265

N

NA
  optical disc formats 1153
NAB (RIAA) curve
  analog discs 1017
Nanometer
  fiber optics 482
Narrow beam loudspeaker system
  message repeaters and evacuation systems 1524
National Association of Broadcasters
  National Association of Broadcastersg 1059
  record discs 1017
National Association of the Deaf (NAD)
  assistive listening systems 1545
National Broadcasting Company
  VU meters 998
National Fire Protection Association
  message repeaters and evacuation systems 1526
National LM381
  integrated circuits 340
Natural convection
  Heatsinks 367
Natural cotton panels
  acoustical treatment for indoor areas 105
Natural sound sources
  computer aided sound system design 1343
Nature of speech
  designing for speech intelligibility 1387
Near field
  acoustics for auditoriums and concert halls 151
Near miss Weber's Law
  psychoacoustics 54
Near-end crosstalk (NEXT)
  wire and cable 414
Near-field loudspeakers
  consoles 918
NEC cable groups
  wire and cable 440
NEC substitution chart
  wire and cable 440
Needed acoustic gain (NAG)
  sound system design 1243
  units of measurement 1659
Negative feedback 709
  amplifier 708
  audio transformers 287
Negative output resistance
  audio transformers 287
Negative-positive-zero (NPO)
  capacitors 257
Neoprene
  wire and cable 406
Network transfer function
  capacitors 253
NFPA 1526

Nickel-Cadmium batteries 697
Nikola Tesla 7
NIOSH
  what's the ear for? 1633
Nits
  display technologies 1579
Noise
  fiber optics 468
  preamplifiers and mixers 737
  resistors 244
  sound system design 1241, 1325
  what's the ear for? 1636
Noise characteristics
  consoles 887
Noise criteria
  acoustical noise control 67
Noise criteria (NC) curves 67
Noise Criteria ratings
  test and measurement 1610
Noise diode
  solid state devices 322
Noise equivalent power (NEP)
  fiber optics 482
Noise figure
  audio transformers 289
  consoles 860
  transistors 331
Noise level analysis
  acoustical noise control 70
Noise reduction
  microphones 514
Noise Reduction Coefficient (NRC)
  acoustical treatment for indoor areas 101
Noise reduction systems 1078
Noise sensing
  designing for speech intelligibility 1392
Noise sources
  consoles 910
  microphones 513
Noise-adaptive threshold automatic microphone mixers 750
Nominal impedance
  computer aided sound system design 1345
  fundamentals of audio and acoustics 33
Nominal voltage
  relays 376
Noninverting 336
Noninverting amplifiers
  amplifiers 338
Noninverting voltage amplifiers 711
Nonlinearity of the ear
  psychoacoustics 49
Non-magnetic materials
  audio transformers 278
Nonmaskable interrupt
  consoles 929
Nonreverberant spaces
  acoustical treatment for indoor areas 114
Nontotalizer
  relays 391
Nordic N9 Standard
  VI meters 1004
Normal link pulses (NLPs)
  digital audio interfacing and networks 1494
Normalized field strength
  assistive listening systems 1547
Normalizing
  filters and equalizers 793
Norris Eyring
  units of measurement 1661
NRZI encoding
  digital audio interfacing and networks 1485
NTSC
  VU meters 1002
NTSC television
  display technologies 1581
Null indicators
  consoles 946
Number of open microphones
  units of measurement 1659
Number of open microphones (NOM)
  preamplifiers and mixers 743, 747
  sound system design 1241
Number of open microphones attenuation (NOMA)
  preamplifiers and mixers 746
Numerical aperture
  fiber optics 460, 482
Numerical diffusers
  acoustical treatment for indoor areas 118
Numerical model
  acoustical modeling and auralization 227–228
Nurse call
  intercoms 1561
Nurse call systems
  intercoms 1568
Nyquist diagram 705, 708
  amplifier 705
Nyquist frequency
  delay 811, 814
Nyquist Sampling Theorem
  DSP technology 1169

O

O type attenuators 775
Oblique mode
  small room acoustics 128
Oblique room modes
  small room acoustics 128
Occluded ear
  personal monitors 1425
Occupational Safety and Health Organization 35
Octave-band equalizers
  sound system design 1286
Off-delay
  relays 391
Ohm
  fundamentals of audio and acoustics 32
Ohm's Law
  fundamentals of audio and acoustics 32
  grounding and interfacing 1181
Ohm's law
  units of measurement 1654
Ohm's law of acoustics
  psychoacoustics 50
OLED
  display technologies 1587
Oliver Heaviside 6
Omnidirectional microphone 493–494
On-delay
  relays 391
One-half wave power supplies 672
One-third-octave band filters

  filters and equalizers 800
Op-amps
  integrated circuits 344, 358
Open circuit
  fundamentals of audio and acoustics 25
Open wire antenna
  assistive listening systems 1552
Open-cell foams
  acoustical treatment for indoor areas 105
Open-circuit stability
  consoles 864
Open-circuit voltage sensitivity
  microphones 516
Opera houses
  acoustics for auditoriums and concert halls 163
Operate time
  relays 376, 383
Operate voltage
  relays 376
Operational amplifier 711, 713
Operational amplifiers 710
  consoles 833, 835
Operational amplifiers-compensating
  consoles 841
Operational transconductance amplifier
  integrated circuits 340
Operational transconductance amplifiers
  integrated circuits 340
Operational voltage amplifiers
  integrated circuits 335
Optic Absorption
  fiber optics 479
Optical cable pull force
  digital audio interfacing and networks 1503
Optical disc formats
  AACS 1152
  Advanced Encryption Standard (AES) 1153
  AGC 1143
  AIFF files 1135
  astigmatism 1138–1139
  audio codecs 1152
  Audio with Video (AV) disc 1149
  auto-focusing 1138
  auto-tracking 1139
  BD 1153
  BD-ROM 1153
  BD-Video 1153
  Blu-ray 1133
  Blu-ray disc system 1153
  CD player hardware architecture 1137
  CD-DA 1133
  CD-R 1133, 1142
  CD-ROM 1133, 1141
  CD-ROM/XA 1142
  CD-RW 1133
  CD-UDF 1143
  CD-WO (Write Once) 1142
  CIRC 1135
  CIRC code 1147
  CIRC decoder 1140
  CIRC encoding algorithm 1136
  CIRC error correction 1141
  coherent-phase laser beam 1138
  Compact Disc Digital Audio 1133
  Compact Disc Read Only Memory 1141
  Compact Disc Recordable (CD-R) 1142
  constant linear velocity (CLV) 1134
  Content Protection for Pre-Recorded Media (CPPM) 1150
  CPPM 1150
  Cross Interleave Reed Solomon Code 1135
  D/A converters 1137
  DAT 1135
  data decoding 1140
  date encoding 1134
  DDP 1134
  decimation filter 1144
  digital signal processing 1140
  Direct Stream Digital 1133, 1144
  Direct Stream Digital (DSD) coding 1144
  Disk Description Protocol 1134
  Dolby Digital 1149
  Dolby Digital (AC-3) 1148
  DSD 1144
  DSP 1140
  DVD 1133
  DVD disc manufacturing 1146
  DVD family 1145
  DVD file format 1146
  DVD formats 1151
  DVD optical disc formats 1145
  DVD+RW 1152
  DVD-Audio 1133, 1145, 1148
  DVD-R 1151
  DVD-RAM 1152
  DVD-ROM 1133, 1145, 1151
  DVD-RW 1152
  DVD-Video 1133, 1145, 1147, 1151
  DVD-Video VAN disc 1147
  EDC/ECC 1141
  EFM encoding 1137
  EFM signal 1140
  EFMPlus modulation 1147
  error concealment algorithms 1140
  finite impulse response 1140
  FIR 1140
  HD DVD 1152
  High Definition DVD 1152
  interpolation 1140
  land 1134
  LPCM 1148
  Meridian Lossless Packing (MLP) 1151
  Microsoft HDi Interactive Format 1153
  Mixed Mode CD 1142
  MLP 1151
  MPEG-1 1149
  MPEG-2 1149
  MPEG-2 AAC 1148
  Multimedia Home Platform (DVB-MHP) 1154
  multiread drives 1143
  NA 1153
  optical pickup 1137
  Optimum Power Calibration (OPC) 1151

  PCM signal 1144
  photodiode array 1140
  pit track 1133
  polycarbonate substrate 1134
  preemphasis 1135
  Reed Solomon Product Code (RS-PC) 1147
  Reed-Solomon decoders 1140
  Reed-Solomon product code (RS-PC) 1141
  reflective data pit surface 1134
  RS-PC 1141
  SACD 1133, 1144
  SDM 1145
  sigma-delta D/A converters 1141
  sigma-delta modulation (SDM) 1145
  signal reconstruction 1140
  SMART 1150
  subcode 1135
  Super Audio CD 1133
  Super Audio Compact Disc (SACD) 1144
  three-beam design pickup 1138
  tracking correction signal 1139
  transversal filter 1140
  UDF 1146
  Universal Disc Format (UDF) Bridge 1146
  Universal Disk Format (CD-UDF) 1143
  video codecs 1152
  WAV files 1135
Optical fiber class
  fiber optics 482
Optical fiber cross section 455
Optical fiber loss
  fiber optics 458
Optical interface
  digital audio interfacing and networks 1483
Optical pickup
  optical disc formats 1137
Optical sSoundtracks
  surroound sound 1593
Optical Time Domain Reflectometer
  fiber optics 462
Optical Time Domain Reflectometer (OTDR)
  fiber optics 462, 482
OpticalCon® fiber optic connector
  fiber optics 473
Optically coupled solid-state relays 390
Optimized method
  gain structure 1229
Optimum Power Calibration (OPC)
  optical disc formats 1151
Optimum source impedance (OSI)
  consoles 861
Opto-coupled silicon-controlled rectifier
  solid state devices 325
Optoelectronic
  fiber optics 482
Optoelectronic Integrated Circuits (OEICs)
  fiber optics 482
Orchestra pit
  acoustics for auditoriums and concert halls 177
Organic Light Emitting Diode (OLED)
  display technologies 1587
Organizations
  International Telegraph and Telephone Consultative Committee (CCITT) 1466
  SMPTE 919
ORTF microphone technique 544
OSHA 35, 1232, 1433
  what's the ear for? 1633
OSHA-permitted noise levels
  what's the ear for? 1634
Ossicles
  personal monitors 1432
OTDR
  fiber optics 462
Outdoor sound
  arraying devices 206
  atmospheric absorption 204
  atmospheric absorption of acoustical energy 204
  atmospheric effects 203
  Bessel array 208
  distributed loudspeaker systems 203, 210
  effects of absorption of sound by air 205
  effects of air temperature 204
  effects of relative humidity 204
  high acoustic pressures 205
  loudspeaker coverage 203
  medium and long throw devices 205
  parabolic reflector 209
  private suite systems 211
  sound projection distances 203
  source level requirements 203
  speed of sound 204
  temperature inversion gradient 204
Outer ear
  personal monitors 1432
  psychoacoustics 43
Output capacitance
  tubes 314
Output capacitance compensation
  integrated circuits 338
Output impedance
  consoles 840
Output transformer
  audio transformer 296
Output transformers
  audio transformers 286, 289–290
Oval window
  psychoacoustics 43
Overdrive
  relays 376
Overdub mode
  consoles 923
Overdubbing
  consoles 828
Overhung coil
  loudspeakers 607
Over-sampling
  delay 812, 814
Oversampling
  consoles 958
Overvoltage protection
  power supplies 681

P

PA loudspeakers
  personal monitors 1415
Packaged loudspeaker systems
  sound system design 1257
Pads

sound system design 1291
PAG values
  personal monitors 1417
Paging
  intercoms 1568
Panel radiators
  loudspeakers 611
Panel resonances
  acoustics for auditoriums and concert halls 184
Panning
  consoles 829
Panpots
  consoles 915
Parabolic effect
  acoustics for auditoriums and concert halls 180
Parabolic microphone 537
Parallel resonant circuit
  filters and equalizers 790
Parallelism
  consoles 961
Parametric equalizer 1449
  consoles 968
Parametric equalizers
  filters and equalizers 801
  sound system design 1287
Parametric room impulse response (PRIR)
  acoustical modeling and auralization 234
Parasitics
  audio transformers 277
Particle of light
  fiber optics 453
Particle velocity
  small room acoustics 127
Party line systems
  intercoms 1561, 1570
Pass band
  filters and equalizers 785
Pass elements
  power supplies 680
Passive crossover networks
  sound system design 1279
Passive device
  sound system design 1275
Passive devices
  sound system design 1297
Passive filters
  filters and equalizers 785–786
Passive mixing
  consoles 914
Passive optical interconnections
  fiber optics 470
Patch editors
  MIDI 1123
Paul Klipsch 5, 12, 17
  loudspeakers 615
Paul W. Klipsch 9, 14
PBX lines
  intercoms 1563
PCM
  delay 813
  digital audio interfacing and networks 1463
PCM signal
  optical disc formats 1144
PDP
  display technologies 1583
Peak amplitude
  gain structure 1225
Peak level detection
  integrated circuits 345
Peak loudness
  fundamentals of audio and acoustics 36
Peak program meter (PPM)
  consoles 920
  VI meters 1003
Peak voltage
  gain structure 1224
Peak-inverse-voltage
  solid state devices 320
Peak-reverse-voltage
  solid state devices 320
Peak-to-peak voltage
  gain structure 1224
Pentagrid
  tubes 311
Pentode
  tubes 311
People absorption
  acoustical treatment for indoor areas 108
Perceived room size
  delay 808
Percentile levels
  acoustical noise control 70
Percussion controllers
  MIDI 1117
Perfect pitch
  psychoacoustics 56
Perforated coverings
  acoustical treatment for indoor areas 105
Perforated membrane absorber
  acoustical treatment for indoor areas 111
Performance characteristics of various microphones 493
Performance data
  loudspeakers 631
Performance zone
  acoustics for auditoriums and concert halls 168
Permalloy®
  audio transformers 283
Permanent threshold shift (PTS)
  what's the ear for? 1633
Permeability
  audio transformers 278
  grounding and interfacing 1189
Persistence
  VI meters 1006
Personal Listening Systems (PLS)
  assistive listening systems 1553
Personal mixing
  personal monitors 1430
Personal monitor mixers
  personal monitors 1430
Personal monitor system 1416
Personal monitors
  ambient microphones 1431
  applications for personal monitors 1425
  auditory nerve 1433
  balanced armature earphones 1424
  belt-pack receiver 1416
  cilia 1433
  cochlea 1433
  comb filtering 1432
  compressors 1431
  custom molded earphones 1423
  custom sleeve earphones 1423
  digital audio devices 1431
  distributed mixing 1430
  drum throne shakers 1430

dual driver earphones 1424
dual mono mixes 1422
dynamic earphones 1424
ear 1432
ear canal 1432
ear drum 1432
ear-buds 1423, 1433
earphone transducers 1424
earphones 1422
effects processing 1431
ethernet (Cat-5) 1430
feedback 1429
FOH 1418, 1427
gain-before-feedback 1417
Hass Effect 1431
hearing conservation 1418
hearing damage 1433
houses of worship 1429
in-ear monitoring 1416, 1420
isolating earphone 1433
latency 1431
limiters 1431
middle ear 1432
mixing for personal monitors 1427
mobility 1418
monitor mixes 1422
mono mixes 1422
MPX 1428
occluded ear 1425
ossicles 1432
outer ear 1432
PA loudspeakers 1415
personal mixing 1430
personal monitor mixers 1430
pinna 1432
portability 1418
potential acoustic gain 1417
reduced vocal strain 1418
safe listening 1432
sound contractors 1429
stapes 1433
stereo image 1429
stereo mixes 1422
stereo monitor mix 1422
stereo multiplexed transmissions 1428
stereo wireless transmission 1428
stereo wireless transmitters 1428
Temporary Threshold Shift (TTS) 1434
tinnitus 1434
Peter D'Antonio 14
Peutz equations
  designing for speech intelligibility 1396
Pew-back distributed system
  sound system design 1270
Phantom power
  MIDI 1111
  preamplifiers and mixers 739
Phantom power for capacitor microphones 510
Pharynx
  psychoacoustics 43
Phase
  cancellation 31
  psychoacoustics 50
  summation 31
  units of measurement 1652
Phase angle
  filters and equalizers 786
  units of measurement 1652
Phase Coherent Cardioid microphone (PCC) 524
Phase comparator
  digital audio interfacing and networks 1478
Phase controlled regulated power supplies 683
Phase correlation
  VI meters 1004
Phase correlation meter
  VI meters 1004
Phase delay
  filters and equalizers 786
Phase difference
  power supplies 675
  units of measurement 1652
Phase distortion 713
  audio transformers 285
  preamplifiers and mixers 738
phase distortion 1077
Phase error
  VI meters 1006
Phase interference
  fundamentals of audio and acoustics 32
Phase interpolation
  computer aided sound system design 1364
Phase locked loop
  digital audio interfacing and networks 1474
Phase margin
  consoles 838
Phase plug
  loudspeakers 601
Phase response
  test and measurement 1613
Phase reversal boxes
  loudspeakers 1350
Phase shift
  consoles 878
  delay 809
  fundamentals of audio and acoustics 33
  units of measurement 1652
Phase spectra
  psychoacoustics 57
Phase spectrum
  psychoacoustics 50
Phase velocity
  units of measurement 1652
Phase-locking
  psychoacoustics 56
Phone plugs
  sound system design 1293
Phonemically Balanced (PB) word lists
  designing for speech intelligibility 1403
Phons
  psychoacoustics 51
Photocells (light-sensitive resistors) 248
Photocurrent
  fiber optics 482
Photodiode
  fiber optics 482
Photodiode array
  optical disc formats 1140
Photon
  fiber optics 482
Photon energy 453
Photophone
  fiber optics 451
Photoresistor

consoles 874
Physical models
  acoustical modeling and auralization 215
Physiological effects of shock current on humans
  wire and cable 443
Pi filter
  power supplies 676
Pi or delta attenuator 774
Piezoelectric
  loudspeakers 604
Piezoelectricity 1030
  analog discs 1030
  loudspeakers 604
Pigtail
  fiber optics 482
PIN photodiode
  fiber optics 467, 482
PIN-FET receiver
  fiber optics 482
Pink noise
  loudspeakers 639
Pinna
  personal monitors 1432
  psychoacoustics 43
Piston directivity
  loudspeakers 634
Piston resistance
  loudspeakers 633
Pit track
  optical disc formats 1133
Pitch
  psychoacoustics 54
Place coding
  psychoacoustics 55
Planar transistor
  solid state devices 326
Planck's law
  fiber optics 453
Plane angle
  units of measurement 1649
Planning fundamentals
  acoustics for auditoriums and concert halls 162
Plasma display
  display technologies 1583
Plasma monitor
  display technologies 1583
Plastic optical fiber (POF)
  fiber optics 482
Plastics
  wire and cable 405
Plate
  tubes 311
Plate dissipation
  tubes 315
Plate efficiency
  tubes 315
Plate resistance
  tubes 314
Platform configuration in concert halls
  acoustics for auditoriums and concert halls 175
Playback
  MIDI 1120
Plenum
  heatsinks 369
Plenum cable
  wire and cable 441
Plenum noise attenuators
  acoustical noise control 92
Plug-ins
  consoles 984
Plug-on transmitter
  microphones 574
Pocket page access
  intercoms 1568
Point source
  fundamentals of audio and acoustics 36
Point source interactions
  loudspeaker clusters 658
Point sources
  computer aided sound system design 1344
Point to point systems
  intercoms 1570
    1561
Polar ETC measurements
  designing for speech intelligibility 1409
PolarFlex microphone system 540
Polarity
  audio transformers 276
  microphones 521
  Test and measurement 1627
  transistors 328
  tubes 313
Polarity testers
  test and measurement 1627
Polarized relay 380
Pole-Zero analysis
  filters and equalizers 796
Pole-zero diagram 706
Poly/Mono mode
  MIDI 1104
Polycarbonate substrate
  optical disc formats 1134
Polyethylene
  wire and cable 405
Polygon ground plan
  acoustics for auditoriums and concert halls 170
Polypropylene
  wire and cable 406
Polyurethane foam panels
  acoustical treatment for indoor areas 106
Pop filters
  microphones 576
Pop protection
  microphones 578
Population inversion
  fiber optics 465
Porous absorbers
  acoustical treatment for indoor areas 102
Porous materials
  Acoustics for auditoriums and concert halls 184
Portability
  personal monitors 1418
Portable and tour sound systems 1317
Positive and negative type power supply 337
Positive feedback
  consoles 838
Positive-intrinsic-negative detector (PIN)
  fiber optics 454
Post-Detection Combining 561
Potential acoustic gain
  personal monitors 1417
  units of measurement 1659
Potential acoustic gain (PAG)
  sound system design 1241
Potential divider

  filters and equalizers 787
Potential Energy (U)
  units of measurement 1649
Potentiometers 250
Potentiometric switching
  consoles 853
Power
  continuous output 1231
  fundamentals of audio and acoustics 25
Power (W)
  units of measurement 1650
Power amplifier 716
  fundamentals of audio and acoustics 25
Power amplifier DSP
  sound system design 1288
Power amplifiers
  sound system design 1290
Power coefficient
  of resistors 245
Power factor (PF)
  capacitors 255
Power isolation transformer 1215
Power Over Ethernet (PoE)
  wire and cable 417
Power over Ethernet (PoE)
  digital audio interfacing and networks 1490
Power Over Ethernet Plus (PoE Plus)
  wire and cable 417
Power ratings
  gain structure 1231
Power ratios 1203
Power sensitivity
  tubes 315
Power supplies
  batteries 692
  battery charging 693
  capacitor filters 674
  charging voltage versus temperature 694
  combination filters 676
  comparator elements 681
  compensation 338
  complex 680
  constant-current regulated 679
  constant-voltage regulated 678
  constant-voltage, constant-current regulated 679
  converters 690
  current limiting 681
  current-regulator circuit 680
  dc to dc converter 725
  diode versus synchronous rectifiers 685
  dynamic line regulation 672
  dynamic load regulation 672
  electromagnetic interference (emi) 682
  equalizing batteries 694
  filters 674
  forward resistance 669
  forward voltage drop 669
  full-wave bridge rectifier 673
  full-wave supplies 672
  high-power regulated supply 682
  inductive filters 675
  inductive reactance 676
  inductive-capacitance filters (LC) 676
  internal output impedance 672
  inverters 690
  magnetic field 676
  maximum peak current 669
  maximum peak inverse voltage 669
  one-half wave supplies 672
  overvoltage protection 681
  pass elements 680
  phase difference 675
  phase-controlled regulated 683
  pi filter 676
  preregulator 681
  rectifier 669
  rectifier circuit chart 670
  reference elements 680
  regulated 677
  regulated supplies 679
  remote sensing 681
  resistance voltage dividers 677
  resistance-capacitance filters 676
  reverse current 669
  reverse resistance 669
  ripple voltage 669
  sampling elements 680
  Schottky diode 686
  single IC, power factor corrected off-line 684
  solid-state zener diode 679
  static line regulation 672
  switch mode 724–725
  switching regulators 682
  synchronous rectification low-voltage 684
  synchronous rectifiers 688
  terminology 669
  thermal regulation 672
  three-phase 674
  ultra capacitor characteristics 691
  ultra capacitor UPS 690
Power supply terminology 669
Power transfer
  consoles 859
  gain structure 1228
  sound system design 1300
Power transformer
  audio transformers 280
Power transformers
  grounding and interfacing 1190, 1194
Power-level meters
  VI meters 1003
Power-line noise
  grounding and interfacing 1193
Power-output meters
  VI meters 1003
Power-sum near-end crosstalk (PS-NEXT)
  wire and cable 415
Preambles
  digital audio interfacing and networks 1465
Preamplifiers
  analog discs 1033
Preamplifiers and mixers
  analog speakerphone 760
  automatic gain control (AGC) 747
  automatic microphone 741
  automatic microphone mixers implemented in software 752
  automatic microphone mixers with matrix mixing to multiple outputs 752
  automatic mixing controller

752
balanced input stage 740
bandwidth 738
build-up of background noise and reverberation 743
capacitance of shielded twisted-pair cable 736
class A circuitry 740
CMRR 735, 740
comb filtering 743
common-mode rejection ratio (CMRR) 739
communication protocols used in automatic mixers 746
conference bridge 758
conference room ceiling systems 761
conference room noise 755
conference room reverberation 756
controls and features of automatic microphone mixers 745
Cooper Sound CS104 four-channel ENG mixer 740
critical distance 757
design objectives for automatic microphone mixers 744
direction dependent automatic microphone mixers 750
echo canceller 760
equivalent input noise (EIN) 738
fixed threshold automatic microphone mixers 748
gain and headroom 737
gain sharing automatic microphone mixers 748
input channel attenuation 747
input impedance 737
Jensen-Hardy Twin-Servo® 990 microphone preamplifier 741
loading loss 735
low-pass LC filter 736
microphone as a signal source 735
microphone mixing 759
microphone preamplifier fundamentals 735
multiple open microphones 742
multivariable dependent automatic microphone mixers 751
noise 737
noise-adaptive threshold automatic microphone mixer 750
number of open microphones (NOM) 743, 747
number of open microphones attenuation 746
phantom power 739
phase distortion 738
real world preamp 739
reduced gain before feedback 743
RF immunity 739
room and building acoustics 755
send level control 760
Shure SCM268 four-channel mixer 740
signal leakage and undesired feed back in a teleconferencing system 759
teleconferencing 754
teleconferencing equipment 758
telephone interface 758
telephone/transmission system interface 757
thermal noise 737
time domain 738
transformers 739
two-wire telephone connection 758
types of automatic microphone mixers 748
variable threshold automatic microphone mixers 748
Precedence effect
  psychoacoustics 60
Precision listening rooms
  Small room acoustics 140
Pre-delays
  consoles 975
Preemphasis
  Optical disc formats 1135
Prefade listen (PFL)
  consoles 827, 918
Preform
  fiber optics 483
Preregulator
  power supplies 681
Pressure
  units of measurement 1650
Pressure open-circuit sensitivity
  computer aided sound system design 1356
Pressure zone microphones (PZM) 522
Pressure-gradient condenser microphone 531
Primary coating
  fiber optics 483
Print through
  tape print thoughg 1072
Private bundles
  digital audio interfacing and networks 1507
Private line
  intercoms 1561
Pro Logic decoding
  surround soound 1596
Processing system
  virtual systems 1454
Processor control
  consoles 855
Processor types
  consoles 960
ProDigi machines
  magnetic recording 1079
Professor Helmholtz 1339
Professor Sabine 1339
Program interrupt (IFB)
  intercoms 1573
Program material
  fundamentals of audio and acoustics 34
  gain structure 1231
Programmable digital signal processor
  DSP technology 1172
Programmable equalizers
  filters and equalizers 802
Programmable gate arrays
  consoles 942
Programmable time delay relay 392
Programs
  Modeler 1265
Progressive delay drivers
  loudspeaker clusters 663

Progressive scan
  display technologies 1581
PROM
  consoles 928
Propagation delay
  consoles 912
Propagation of sound waves
  acoustical modeling and auralization 217
Propagation velocity
  grounding and interfacing 1184
Properties of plastic insulation
  wire and cable 404
Properties of rubber insulation
  wire and cable 404
Proportional Point Source
  loudspeakers 614
Proprietary acoustical doors
  acoustical noise control 83
Proprietary acoustical windows
  acoustical noise control 86
Proprietary silencers
  acoustical noise control 93
Protecting concert-goers
  what's the ear for? 1636
Protecting hearing
  what's the ear for? 1635
Protecting loudspeakers
  sound system design 1276
Protecting the audience
  what's the ear for? 1642
Protecting the community
  what's the ear for? 1636
Protection mechanisms 719
  automatic turn-on muting 720
  capacitor high pass filter 720
  fuse 719
  thermal sensor 720
  transformer 720
Protocol
  digital audio interfacing and networks 1505
Proximity effects
  microphones 497
Psychoacoustics
  absolute pitch 56
  acoustic reflex 45
  amplitude spectra 57
  amplitude spectrum 50
  audio encoding 48
auditory area 50
auditory filters 49
auditory system 45
auricle 43
A-weighting level 52
backward masking 48
Bark scale 49
binaural 57
binaural hearing 60
binaural pitches 57
binaural process 56
bone conduction 45
B-weighting level 52
cent 55
chords 55
cochlea 43
cocktail party effect 60
combination tones 50
complex tone 54
compressed function 52
concha 43
cone of confusion 58–59
contralateral ear 58
critical bands 49
critical bandwidths of the human ear 49
critical distance 61
C-weighting level 52
decoder 48
diffraction of the head 58
direct-to-reverberant ratio 61
distance cues 60
distance perception
  psychoacoustics 60
dynamic changes 54
dynamic range 51
ear canal (auditory meatus) 43
ear molds 59
eardrum 43
equal loudness contours 51
equal temperament 55
equivalent rectangular bandwidth (ERB) 49
ERB number 49
eustachian tube 43
external ear 43
externalization 59
Fletcher-Munson curves 51
forward masking 48
Fourier analyzer 51
Franssen effect 60
frequency difference limen 56
frequency range 51
frequency selectivity 52
Haas effect 60
hair cells 43
harmonic distortion 50
head shadow 58
head-related transfer function (HRTF) 58
hearing damage 45, 52
hearing loss 47
hearing threshold 51
hydrophone 45
incus 43
individual differences 43
inner ear 43, 45
integration times 54
interaural intensity difference (IID) 57
interaural level difference (ILD) 57
interaural phase difference 50
interaural phase difference (IPD) 57
interaural temporal difference 58
interaural time difference (ITD) 57
just-noticeable difference (JND) 54
Just-tuning 55
LEDR (Listening Environment Diagnostic Recording) 59
localization cues 57
logarithmic scale 52, 55
lossless encoders 48
loudness 51–52
loudness versus bandwidth 53
malleus 43
masker 48
masking 48
mechanical system of the middle ear 46
mel 55
melodic intervals 55
melody 55–56
microphone 45
middle ear 43, 45
middle ear infection 45

minimum audible 51
near miss Weber's Law 54
nonlinearity of the ear 49
Ohm's law of acoustics 50
otoacoustic emissions 50
ototoxic medicines 47
outer ear 43
oval window 43
perfect pitch 56
pharynx 43
phase 50
phase spectra 57
phase spectrum 50
phase-locking 50, 56
phons 51
pinna 43
pitch 54
place coding 55
precedence effect 60
psychophysical experiments 43
psychophysical tuning curve 48
pure tone 54
Pythagorean scale 55
range of pitch 56
resonant frequency 45
reverberator 59
RMS detectors 54
root note 55
round window 43
sagittal planes 58
shape of the auditory filters 49
signal detection 60
simultaneous masking 48
sones 53
sound color 57
sound pressure level (SPL) 50
spacial hearing 57
spatial filtering 60
spatial-filtering 44
spectrogram 51, 57
spherical head model 44, 58
stapes 43
statistical methods 43
Steven's rule 56
temporal bone 44
temporal coding 55–56
threshold of hearing 50
thresholds 50
timbre 57
time window 51
tinnitus 50
tympanic membrane 43
virtual pitch 57
wavelet analysis 51
Weber fraction in dB 54
Weber's Law 54
white noise 57

Psychophysical tuning curve
  psychoacoustics 48
Pulse code modulation
  digital audio interfacing and networks 1463
Pulsed Code Modulation
  delay 813
Punch through
  transistors 332
Pure tone
  psychoacoustics 54
Pyramid tracing
  computer aided sound system design 1372
Pythagorean scale
  psychoacoustics 55
PZM directivity
  microphones 524
PZM direct-to-reverberant ratio effects
  microphones 525
PZM frequency response anomalies caused by boundaries
  microphones 526
PZM frequency-response effects
  microphones 525
PZM sensitivity effects
  microphones 524

Q

Q
  consoles 879, 883, 888
  inductors 267
Q factor
  acoustical treatment for indoor areas 110
Quadratic residue diffuser (QRD)
  acoustical treatment for indoor areas 118
Quality factor (Q
  capacitors) 255
Quality of Service (QoS)
  digital audio interfacing and networks 1496
Quantization
  DSP technology 1170
  sampling 813
Quantization error
  DSP technology 1171
Quantization noise
  delay 813–814
Quantum efficiency
  fiber optics 468
Quarter-frame messages
  MIDI 1129
Quick-operate relay 380

R

R. A. Greiner 13
Radial (or sectoral) horns
  loudspeakers 612
Radiation
  heatsinks 368, 370
Radiator types
  loudspeakers 609
Radio frequencies
  fundamentals of audio and acoustics 28
Radio frequency spectrum
  units of measurement 1654
Radio guide designations (RG)
  wire and cable 424
Radio-frequency inductors
  consoles 912
Raised-slab or jack-up system
  acoustical noise control 81
RAM
  delay 811, 814
RAMSETE
  computer aided sound system design
  1372
Random access memory
  delay 811
Random diffusers
  acoustical treatment for indoor areas 119
Range of pitch
  psychoacoustics 56
Range of wireless transmission
  microphones 566
RASTI
  acoustics for auditoriums and concert halls 159
  designing for speech intelligibil-

ity 1405
RaSTI
  designing for speech intelligibility 1406–1407
Rate of decay
  small room acoustics 129
Rate of delivery
  designing for speech intelligibility 1401
Rate of electron flow
  fundamentals of audio and acoustics 32
Rated impedance
  fundamentals of audio and acoustics 33
Rated sensitivity
  computer aided sound system design 1346
Ray
  fiber optics 483
Ray radiation
  computer aided sound system design 1372
Ray tracing
  computer aided sound system design 1367, 1371
Rayleigh scattering
  fiber optics 459, 483
Ray-tracing models
  acoustical modeling and auralization 223
RCA-type phono plugs
  sound system design 1293
R-DAT
  magnetic recording 1080
Reactance
  capacitive 33, 270
  chart 271
  consoles 878
  fundamentals of audio and acoustics 33
  inductive 33, 1548
  Inductive reactance 1183
  inductive reactance 248, 258, 265–267, 270, 864, 880, 891
Reactive ground effects
  consoles 850
Read only memory
  read only memory 1050
Read only memory (ROM)
  message repeaters and evacuation systems 1528
RealAudio
  digital audio interfacing and networks 1461
Realtime analyzers
  loudspeakers 639
Real-time auralization
  acoustical modeling and auralization 233
Real-time messages
  MIDI 1127
Real-time programming
  DSP technology 1174
Real-world preamp
  preamplifiers and mixers 739
Rear and surround cluster loudspeaker systems
  sound system design 1262
Recall display
  consoles 945
Receiver
  acoustical modeling and auralization 218
  fiber optics 483
  sensitivity
    fiber optics 483
  specifications
    fiber optics 476
Receiver quality
  wire and cable 423
Receivers 1555
Reception modes
  MIDI 1104
Reconstruction filters
  consoles 958
Record care
  analog discs 1036
  brushes 1036
Record Head 1061
Record Industry Association of America
  analog discs 1017
Record-cleaning machines
  analog discs 1036
Recording
  MIDI 1119
Recording characteristics
  analog discs 1017
Recording mode
  consoles 923
Rectifier
  power supplies 669
Rectifier circuit chart
  power supplies 670
Rectifiers
  light-activated silicon controlled 325
  opto-coupled silicon-controlled rectifier 325
  SCR (silicon-controlled 324
  silicon 320
  silicon-controlled rectifier 322
  switching diodes 321
  thyristors 322
Recursive processing
  consoles 965
Reduced gain before feedback
  preamplifiers and mixers 743
Reduced vocal strain
  personal monitors 1418
Reducing noise pickup
  sound system design 1306
Reed relay packages
  relays 383
Reed relays 381
Reed Solomon Product Code (RS-PC)
  optical disc formats 1147
Reed-Solomon decoders
  optical disc formats 1140
Reed-Solomon product code (RS-PC)
  optical disc formats 1141
Reference elements
  power supplies 680
Reference levels
  VU meters 1000
Reference point array
  loudspeaker clusters 653
Reflected wave
  fundamentals of audio and acoustics 29
Reflection
  acoustical treatment for indoor areas 120
  fiber optics 454, 483
  sound system design 1245
Reflection filter

  microphones 578
Reflection free measurements
  test and measurement 1614
Reflection free zone
  small room acoustics 141
Reflection manipulation
  small room acoustics 138
Reflective data pit surface
  optical disc formats 1134
Reflectogram
  acoustical modeling and auralization 221
  computer aided sound system design 1376
Refraction
  acoustical treatment for indoor areas 120
  fiber optics 454, 483
Refractive index
  fiber optics 455, 483
Register balance measure BR
  acoustics for auditoriums and concert halls 162
Regulated power supplies 677, 679
Related crosstalk
  consoles 919
Relative loudness
  fundamentals of audio and acoustics 36
  VI meters 1006
Relative phase
  fundamentals of audio and acoustics 31
Relay layouts 374
Relay loads 378
Relays
  ac 380
  ac differential 381
  bifurcated contacts 377
  bistable magnetic latching 380
  bounce 383–384
  capacitive loads 379
  consoles 850
  contact arrangements 378
  contact bounce 378
  contact characteristics 377
  contact current 377
  contact resistance 383–384
  contact resistance stability (CRS) 381
  darlington circuit 390
  dc 380
  dc loads 379
  differential polar 381
  DIP 374
  direct-coupled solid state 390
  dry reed 381–382
  dry reed switches 387
  dual-coil latching 380
  dv/dt transients 389
  effects of various loads 379
  electromagnetic interference (emi) 379
  electromechanical 379
  environmental temperature effects 382
  glossary of terms 374
  high peak transient-voltage protection 389
  high speed switching 386
  high transient voltages 377
  high-surge current protection 389
  inductive loads 379
  insertion loss 385
  isolation 385
  j-bend 386
  latching 376, 380
  layout 374
  leadless 386
  loads 378
  low load current protection 390
  low-level logic switching 382
  magnetic field 380–381
  magnetic interaction 381
  make-or-break current 377
  marginal 380
  MEMS switches 387
  mercury-wetted contact 381, 384
  mercury-wetted reed 384
  metal-oxide varistors (MOV) 389
  momentary actuation 392
  motor loads 379
  Mu-metal 382
  nontotalizer 391
  on-delay 391
  operate time 383
  optically coupled solid-state 390
  polarized 380
  programmable time delay 392
  quick-operate 380
  reed 381
  reed packages 383
  release time 383
  return loss 385
  rf 384
  rf coaxial shielded 384
  rise time 385
  sensitivity 383
  short-circuit protection 389
  SIP 4 Pin 374
  skin effect 386
  slow operate 380
  solid-state time-delay 391
  static contact resistance (SCR) 381
  surface mount 386
  totalizer/preset counter 391
  transformer-coupled solid-state 390
  types of time-delay relays 391
  voltage standing wave ratio 385
Release time
  consoles 901
  relays 376, 383
Release voltage
  relays 376
Remanence
  remanencerecording 1066
Remote control pulses
  digital audio interfacing and networks 1477
Remote Mic-Kill
  intercoms 1575
Remote sensing
  power supplies 681
Repeat coil
  audio transformers 300
Repeaters
  fiber optics 469, 483
Reservation packets
  digital audio interfacing and networks 1505
Resetting functions
  consoles 946
Residential intercom systems

- intercoms 1564
- Residential music systems
  - loudspeakers 618
- Residual field
  - audio transformers 279
- Resistance 243
  - audio transformers 280
  - units of measurement 1650
  - winding resistance 250
  - wire and cable 397
- Resistance temperature detectors 249
- Resistance voltage dividers
  - power supplies 677
- Resistance-capacitance filters
  - power supplies 676
- Resistive pads
  - gain structure 1229
- Resistivity 247
- Resistor networks 248
- Resistors 243
  - ac resistance 245
  - ambient temperature effect 245
  - carbon-composition 247
  - carbon-film 247
  - characteristics 243
  - color codes 244
  - contact resistance 250
  - forced air cooling 246
  - grouping 245
  - hot-spot temperature 244
  - in parallel 247
  - in series 246
  - inductive reactance 248
  - metal film 247
  - networks 248
  - noise 244
  - non-inductive 248
  - photocells (light-sensitive resistors) 248
  - potentiometers 250
  - power coefficient 245
  - pulse operation 246
  - resistance 243
  - resistance temperature detectors 249
  - resistivity 247
  - resistor networks 248
  - rheostats 250
  - temperature coefficient of resistance (TCR) 243
  - thermistors (thermal-sensitive resistors) 249
  - thermocouples 249
  - thin-film 248
  - types of resistors 247
  - variable 248
  - varistors (voltage-sensitive resistors) 249
  - voltage coefficient 243
  - winding 248
  - winding resistance 250
  - wirewound 248
  - wirewound potentiometers 250
- Resolution
  - consoles 954
  - display technologies 1580, 1589
- Resolvers
  - consoles 947
- Resonance
  - consoles 879
- Resonant absorbers
  - acoustical treatment for indoor areas 109
- Resonant frequency
  - inductors and capacitors 270
  - psychoacoustics 45
- Resonant Q
  - consoles 879
- Responsivity
  - fiber optics 483
- Retail store applications
  - message repeaters and evacuation systems 1517
- Retentivity
  - retentivitycording 1066
- Reticulating path
  - delay 809
- Return loss
  - relays 376, 385
  - wire and cable 420
- Return loss (RL)
  - wire and cable 415
- Reverberance measure H
  - acoustics for auditoriums and concert halls 161
- Reverberant field
  - test and measurement 1620
- Reverberant room problem
  - sound system design 1322
- Reverberation
  - consoles 828, 962
  - delay 808
  - designing for speech intelligibility 1394
  - loudspeaker clusters 655
  - small room acoustics 135
  - sound system design 1245
- Reverberation chamber
  - acoustical treatment for indoor areas 98
- Reverberation directivity index
  - computer aided sound system design 1349
- Reverberation on Demand System (RODS)
  - acoustics for auditoriums and concert halls 193
- Reverberation radius
  - acoustics for auditoriums and concert halls 151
- Reverberation synthesis
  - delay 808
- Reverberation time
  - acoustical treatment for indoor areas 114
  - test and measurement 1611
- Reverberation time (RT60)
  - acoustics for auditoriums and concert halls 148, 165
  - designing for speech intelligibility 1393, 1396
- Reverberation time duration
  - acoustics for auditoriums and concert halls 148
- Reverberator
  - psychoacoustics 59
- Reverse current
  - power supplies 669
- Reverse recovery time
  - solid state devices 324
- Reverse resistance
  - power supplies 669
- Reverse voltage
  - of a capacitor 259
- Reverse-blocking thyristor
  - solid state devices 322
- RF coaxial shielded relay 384

RF immunity
  preamplifiers and mixers 739
RF interference
  grounding and interfacing 1203, 1210
RF power output of wireless microphones 560
RF relays 384
RF signal strength
  microphones 568
RF systems
  message repeaters and evacuation systems 1520
RFI
  digital audio interfacing and networks 1503
RGB cable 421
Rheostats 250
RIAA curve
  analog discs 1017
Ribbon loudspeakers
  loudspeakers 606
Richard C. Heyser 17
Richard Heyser 1616
Rifle microphone 537
Rigging
  sound system design 1312
Right-hand-rule
  inductors 268
Rigid disks
  magnetic recording 1084
Ring core
  audio transformers 280
Ring radiator
  loudspeakers 611
Ripple voltage
  power supplies 669
Rise time
  relays 385
rms detection
  consoles 908
RMS detectors
  psychoacoustics 54
RMS level detection
  integrated circuits 345
Robert Millikan 7
Robert Stephens 11
Room acoustic simulation
  computer aided sound system design 1370
Room acoustics
  acoustics for auditoriums and concert halls 190
Room and building acoustics
  preamplifiers and mixers 755
Room constant
  sound system design 1247
Room evaluation
  sound system design 1262
Room impulse response
  computer aided sound system design 1340
Room modes
  small room acoustics 127
Room shape
  acoustics for auditoriums and concert halls 167
  small room acoustics 138
Room volume
  acoustics for auditoriums and concert halls 165
Room-Acoustic Analysis Module
  computer aided sound system design 1373
Room-acoustic measures
  computer aided sound system design 1371
Room-acoustical criteria
  acoustics for auditoriums and concert halls 147
Room-acoustical planning
  acoustics for auditoriums and concert halls 164
Rooms for entertainment
  small room acoustics 142
Rooms for microphones
  Small room acoustics 141
Rooms with variable acoustics
  acoustics for auditoriums and concert halls 164
Room-within-a-room
  acoustical noise control 82
Root note
  psychoacoustics 55
Root-mean-square (rms) voltage
  gain structure 1224
Rotary digital audio tape
  magnetic recording 1080
Round window
  psychoacoustics 43
Router
  digital audio interfacing and networks 1496
Router mixing
  consoles 993
Routing
  consoles 850, 952
  digital audio interfacing and networks 1496
RS 232 1448
RS 485 1448
RS-232 746
RS232 1448
RS-422 746
RS485 1448
RS-PC
  optical disc formats 1141
RT equations
  acoustical treatment for indoor areas 114
Rubber
  wire and cable 406
Rudy Bozak 5

S

Sabine
  small room acoustics 135
  units of measurement 1661
Sabine absorption coefficient
  acoustical treatment for indoor areas 101
Sabine Phantom Mic Rider
  microphones 573
SACD
  optical disc formats 1133, 1144
Safe listening
  personal monitors 1432
Safety grounding
  grounding and interfacing 1190–1191
Sagittal planes
  psychoacoustics 58
Sallen-Key filter
  consoles 888
Sallen-Key filters
  filters and equalizers 794
Sample and Hold
  delay 812
Sample frequencies
  digital audio interfacing and net-

works 1485
Sample Rate Converter
virtual systems 1443
Sample Rate Converter (SRC)
Digital audio interfacing and networks 1463
Sampler
MIDI 1115
Sample-Rate Convertors (SRCs)
consoles 958, 976
Sampling
delay 811, 813
Sampling elements
power supplies 680
Sampling rate
consoles 955
DSP technology 1171
SASS microphones 550
Satellite links
intercoms 1572
Saturate
audio transformers 278
Saul Marantz 5
SC fiber optic connector 472
Scale model test
computer aided sound system design 1340
Scaled model simulation
computer aided sound system design 1341
Scaling
display technologies 1581
filters and equalizers 793
Scan conversion
display technologies 1581
Scan monitoring
intercoms 1568
Scanning
display technologies 1580
Scattering
acoustical treatment for indoor areas 117
fiber optics 459
Scattering coefficient s
acoustics for auditoriums and concert halls 180
Scattering coefficients
acoustical treatment for indoor areas 118
Scattering data
computer aided sound system design 1354
School applications
message repeaters and evacuation systems 1518
Schottky diode
power supplies 686
Schroeder diffusers
acoustical treatment for indoor areas 115
SCR
solid state devices 322, 324
Screen grid
tubes 311, 315
SDM
delay 814
optical disc formats 1145
Sealed box system
loudspeakers 635
Second order filter
filters and equalizers 788
Second-order filters
consoles 890
Security audio monitoring
intercoms 1564
Security intercom systems
intercoms 1563
Selenium rectifiers
solid state devices 320
Semiconductor pickup cartridges
analog discs 1029
Semiconductors
LM381 340
solid-state devices 317
Semi-parametric equalizers
filters and equalizers 801
Send level control
preamplifiers and mixers 760
Sensaphonics Hearing Conservation 1435
Sensitivity
loudspeakers 627
microphones 516
relays 383
Sensitivity level
computer aided sound system design 1345
Sensitivity of the eye
assistive listening systems 1554
Sequencers
MIDI 1118
Sequences
DSP technology 1162
Serial audio format
consoles 976
Serial digital video (SDI)
wire and cable 423
Serial networking
consoles 992
Series leakage inductance
audio transformers 286
Series resonant circuit
filters and equalizers 790
Servo-Drive loudspeaker 606
Servos
magnetic recording 1086
SFF fiber optic connector 472
Shadowing
computer aided sound system design 1367
Shape of auditory filters
psychoacoustics 49
Shelving equalizer
consoles 969
Shield, coaxial
relays 376
Shield, electrostatic
relays 376
Shield, magnetic
relays 377
Shield-current induced noise (SCIN)
grounding and interfacing 1201
Shielded power cable
wire and cable 444
Shielded twisted pair (STP)
digital audio interfacing and networks 1470
Shielding 444
grounding and interfacing 1187–1188
sound system design 1301
wire and cable 425
Shielding to reduce noise pickup
sound system design 1306
Shields
wire and cable 426
Shock mounts
microphones 578

Shockley breakover diode
  solid state devices 322
Short interval reflections
  test and measurement 1609
Short-circuit protection
  relays 389
Shotgun microphone 532
Show theaters
  acoustics for auditoriums and concert halls 164
Shure Brothers 12
Shure SCM268 four-channel mixer
  preamplifiers and mixers 740
SI system
  units of measurement 1647
Sidechain sources
  consoles 909
Sidechain time constants
  consoles 900
Sidney N. Shure 12
Sight lines
  acoustics for auditoriums and concert halls 173
Sigma Delta Modulator
  delay 814
Sigma-delta conversion
  consoles 957
Sigma-Delta converter
  DSP technology 1171
Sigma-delta convertors
  consoles 956
Sigma-delta D/A converters
  optical disc formats 1141
Sigma-delta DACs
  consoles 958
Sigma-Delta Modulation
  delay 814
Sigma-delta modulation (SDM)
  optical disc formats 1145
Sigma-Delta Modulator
  delay 814
Signal alignment
  sound system design 1276
Signal alignment in cluster design
  sound system design 1268
Signal chain
  fundamentals of audio and acoustics 25
Signal delay
  sound system design 1307
  units of measurement 1662
Signal delay for an under-balcony distributed system
  sound system design 1307
Signal delay for loudspeaker clusters
  sound system design 1307
Signal detection
  psychoacoustics 60
Signal grounding
  grounding and interfacing 1192
Signal leakage and undesired feed back in a teleconference system
  preamplifiers and mixers 759
Signal Loss
  fiber optics 458
Signal processing components
  sound system design 1285
Signal Processing Control 980
Signal quality
  Grounding and interfacing 1212
Signal reconstruction
  optical disc formats 1140
Signal rise time
  relays 376
Signal switching
  consoles 850
Signal to noise ratio
  designing for speech intelligibility 1390
  DSP technology 1171
Signal to noise ratio (SNR)
  units of measurement 1660
Signal transmission characteristics
  audio transformers 302
Signal-level compatibility
  sound system design 1298
Signal-to-noise ratio
  gain structure 1227
Signal-to-noise ratio (SNR)
  analog discs 1016
SII Speech Intelligibility Index
  designing for speech intelligibility 1409
Silencers
  acoustical noise control 93
Silicon
  solid state devices 318
Silicon rectifiers
  solid state devices 320
Silicone
  wire and cable 406
Simple active filter 794
Simple Network Management Protocol (SNMP)
  digital audio interfacing and networks 1495
Simulated resonance
  consoles 883
Simulation of complex loudspeakers
  computer aided sound system design 1360
Simulation of microphones
  computer aided sound system design 1369
Simulation of the human head
  computer aided sound system design 1369
Simultaneous masking
  psychoacoustics 48
Sine wave
  fundamentals of audio and acoustics 30
Sine wave power
  gain structure 1231
Single chip DLP technology
  display technologies 1586
Single conductor wire
  wire and cable 406
Single ended bootstrap
  integrated circuits 358
Single entry cardioid microphone 499
Single IC, power factor corrected, off-line power supply 684
Single mode fiber
  fiber optics 456, 483
Single number STC rating
  acoustical noise control 73
Single optical fiber
  fiber optics 454
Single-channel console
  consoles 948
Single-ended power amplifier
  audio transformers 294
Single-mode fiber optic cable
  digital audio interfacing and networks 1490, 1502

Single-order networks
  consoles 876
Sinusoidal waveform
  gain structure 1227
Sir Humphry Davy 5
Sir Isaac Newton 38
Site noise survey
  acoustical noise control 69
Site selection
  acoustical noise control 69
Skating force
  analog discs 1023
Skin depth
  wire and cable 402
Skin effect
  inductors 268
  relays 386
  wire and cable 402
SLARM™
  what's the ear for? 1638
SLARMSolution™
  what's the ear for? 1637
SLARMSolution™ applications
  what's the ear for? 1641
Slew rate
  grounding and interfacing 1187
Slew-rate effects
  consoles 844
Slew-rate limitations
  consoles 836
SLM
  test and measurement 1609
Sloping of tiers
  acoustics for auditoriums and concert halls 173
Slow-operate relays 380
SMA connector 471
Small room acoustics
  anechoic chamber 141
  angle of incidence 139
  axial mode 128
  Bolt footprint 133
  Bonello criteria 132
  comb filters 139
  control room 140
  damping factor 131
  Energy Time Curve 136
  Envelope Time Curve 136
  LEDE 141
  Lord Rayleigh 127
  mean free path 135
  MFP 135
  modal bandwidth 131
  modal distribution 131
  modal potency 131
  modal room resonances 128
  modes in nonrectangular rooms 133
  oblique mode 128
  oblique room modes 128
  particle velocity 127
  precision listening rooms 140
  rate of decay 129
  reflection manipulation 138
  reflective free zone 141
  reverberation 135
  room modes 127
  room shape 138
  rooms for entertainment 142
  rooms for microphones 141
  Sabine 135
  small room design factors 140
  summation of modal effect 134
  tangential mode 128
  tangential room modes 128
Small room design factors
  small room acoustics 140
Small room models
  acoustical modeling and auralization 228
SMART
  optical disc formats 1150
Smooth curved surfaces
  acoustics for auditoriums and concert halls 179
SMPTE
  Audio transformers 284
  MIDI 1128
  VI meters 1005
SMPTE serial digital performance specifications
  wire and cable 424
SMPTE/MTC conversion
  MIDI 1129
SMPTE-to-MTC converter
  MIDI 1128
Snake cable
  wire and cable 408
Snell's Law
  fiber optics 455
SNR
  analog discs 1033
  delay 810, 814
  DSP technology 1171
  gain structure 1226–1227
  SNRnetic recording 1059
Soft failure
  relays 377
Software sequencers
  MIDI 1119
Solid angle (A)
  units of measurement 1649
Solid state devices
  alloy-junction transistor 325
  alpha 329
  avalanche diode 321
  avalanche region 320
  beta 329
  breakdown region 320
  diacs 323, 325
  diodes 320
  doping agent 318
  double-diffused epitaxial mesa transistor 326
  drift-field transistor 325
  field effect transistor 326
  germanium 317
  germanium crystals 319
  grown-junction transistor 325
  insulated-gate transistor 327
  LED 325
  light activated silicon controlled 325
  mesa transistors 325
  microalloy diffused transistor 325
  MOSFET transistor 327
  noise diodes 322
  n-type germanium 318
  opto-coupled solid state silicon controller rectifier 325
  peak-inverse-voltage 320
  peak-reverse-voltage 320
  planar transistor 326
  reverse blocking thyristor 322
  reverse-recovery time 324
  SCR 322, 324
  selenium rectifiers 320
  semiconductors 317
  Shockley breakover diode 322

  silicon 318
  silicon controlled rectifier 322, 324
  silicon rectifiers 320
  small signal diode 321
  step-down transformer 331
  switching diodes 321
  thyristors 322
  transistor bias circuits 329
  transistor current flow 328
  transistor equivalent circuits 328
  transistor forward-current-transfer ratio 329
  transistor gain-bandwidth product 329
  transistor input resistance 328
  transistor internal capacitance 331
  transistor noise figure 331
  transistor polarity 328
  transistor punch through 332
  transistor small and large signal characteristics 330
  transistors 325
  triacs 323, 325
  tunnel diode 322
  turnoff time 324
  turnon time 324
  varactor diodes 322
  zener diode 321
Solid-state time-delay relays 391
Solid-state zener diode
  power supplies 679
Solo, Solo-free
  consoles 918
Sones
  psychoacoustics 53
SONEX
  acoustical treatment for indoor areas 106
Sony Dynamic Digital Sooound (SDDS)
  surround sooound 1595
Sound
  absorption from trees and shrubs 230
  articulation loss of consonants 1396, 1404
Sound absorbers
  acoustics for auditoriums and concert halls 183
Sound Absorption Average (SAA)
  acoustical treatment for indoor areas 101
Sound absorption coefficient
  acoustics for auditoriums and concert halls 149, 183
Sound barrier
  acoustical noise control 71
Sound color
  psychoacoustics 57
Sound coloration measures
  acoustics for auditoriums and concert halls 160
Sound contractors
  personal monitors 1429
Sound delay systems
  acoustics for auditoriums and concert halls 191
Sound design
  computer aided sound system design 1374
Sound energy component
  acoustics for auditoriums and concert halls 152
Sound field components
  designing for speech intelligibility 1394
Sound field structure
  computer aided sound system design 1339
Sound form
  acoustics for auditoriums and concert halls 165
Sound intensity
  units of measurement 1651
Sound level measurements
  Test and measurement 1608
  test and measurement 1609
Sound level meter
  acoustical noise control 67, 70
  test and measurement 1608
sound level meter 1608
Sound lock corridor
  acoustical noise control 83
Sound masking system
  sound system design 1330
Sound pressure level
  acoustics for auditoriums and concert halls 153
  loudspeakers 627
  test and measurement 1608
Sound pressure level (SPL)
  psychoacoustics 50
  units of measurement 1657, 1660
Sound pressure measurements
  test and measurement 1608
Sound production
  loudspeakers 597
Sound propagation
  fundamentals of audio and acoustics 36
Sound reflection coefficient
  acoustics for auditoriums and concert halls 183
Sound reproduction
  loudspeakers 597
Sound system design
  70.7 V loudspeaker transformer 1291
  70.7 V or 100 V loudspeaker system 1291
  ac receptacles 1302
  acoustic gain 1240
  active balanced inputs 1305
  aiming loudspeakers 1263
  Alcons 1251, 1254, 1263, 1265–1267, 1272, 1274
  artificial ambience 1328
  attenuation of sound indoors 1248
  attenuation with increasing distance 1240
  audio and video teleconferencing 1329
  audio signal transmission via computer networks 1333
  automatic equalization 1312
  automatic microphone mixing 1282
  automatic mixing problems 1285
  balanced line 1296
  barrier block connectors 1293
  cable 1292
  capacitors 1278
  Cat5 connectors 1293
  ceiling loudspeaker systems

1270
central cluster loudspeaker 1260
central cluster plus distributed system 1269
CobraNet 1290
commissioning 1292
compression drivers and horns 1256
compressors 1286
computer aided measurement systems 1334
computer aided system design 1333
cone loudspeaker enclosures 1255
cone loudspeakers 1255
conference room sound system 1328
connectors 1293
constant Q equalizers 1287
critical distance 1247
crossover networks 1275
dBu 1299
dBV 1300
definitions 1240
delay 1288
designing a complex cluster 1265
designing a distributed ceiling system 1271, 1273
digital audio connectors 1296
digital audio networking 1290
digital echo canceller 1329
digital signal processing 1288
direct sound 1264
direct/reverberant ratio 1246
distributed ceiling loudspeaker systems 1269
distributed column system 1270
distributed loudspeaker system 1268
distributed systems outdoors 1323
DSP equalizers 1287
echo 1245
echoes outdoors 1323
effect of directional loudspeakers 1253
effect of directional microphones 1243, 1253
effect of humidity 1244
effect of temperature layers 1244
effect of wind 1244, 1326
electrical power required (EPR) 1241
electronic crossovers 1288
equalization 1307
equalization test setup 1309
equalizers 1286
equivalent acoustic distance (EAD) 1242
ethernet-style connector 1296
Euro-Block connectors 1293
evaluating loudspeaker sound quality 1259
feedback 1241, 1264
feedback stability margin (FSM) 1241
fuses 1278
geometrically complex room 1253
graphic equalizers 1286
ground loops 1304
grounding 1301
grounding for safety outdoors 1303
grounding to reduce external noise pickup 1304
headroom 1241, 1275
high-frequency attenuation in air 1322
high-pass filters 1288
human ear 1260
impedance compatibility 1297
impedance matching 1297
indoor sound reinforcement system 1245
inverse square law 1240
left-center-right cluster loudspeaker system 1261
limiters 1279, 1286
line array loudspeaker systems 1257
loudspeaker components 1255
loudspeaker connectors 1295
loudspeaker failure modes 1276
loudspeaker pest damage 1280
loudspeaker power capacity specifications 1277
loudspeaker processing DSP 1288
loudspeaker protection devices 1278
loudspeaker Q 1246
loudspeaker systems 1255, 1260
loudspeakers for distributed systems 1270
low rt60 rooms 1254
low-pass filters 1288
manual/graphical cluster design tools 1264
matrix mixing 1285
microphone snake cables 1295
mix groups, Auxiliary groups and matrix mixing 1281
mixers 1280
mixing consoles 1280
mix-minus mixer 1285
multicluster system 1270
multi-function digital signal processing system 1288
needed acoustic gain (NAG) 1243
noise 1241, 1325
number of open microphones (NOM) 1241
octave-band equalizers 1286
packaged loudspeaker systems 1257
pads 1291
parametric equalizers 1287
passive crossover networks 1279
passive device 1275
passive devices 1297
pew-back distributed system 1270
phone plugs 1293
portable and tour systems 1317
potential acoustic gain (PAG) 1241
power amplifier DSP 1288
power amplifiers 1290
power transfer 1300
protecting loudspeakers 1276

protecting loudspeakers against weather 1280
RCA-type phone plugs 1293
rear and surround cluster loudspeaker system 1262
reducing noise pickup 1306
reflection 1245
reverberant room problem 1322
reverberation 1245
rigging 1312
room constant 1247
room evaluation 1262
shielding 1301
shielding to reduce noise pickup 1306
signal alignment 1276
signal alignment in cluster design 1268
signal delay 1307
signal delay for loudspeaker clusters 1307
signal delay for under-balcony distributed system 1307
signal processing components 1285
signal-level compatibility 1298
sound masking system 1330
Source Independent Measurement (SIM) test equipment 1312
specifications for sound reinforcement 1280
split cluster loudspeaker system 1261
sports stadiums and outdoor systems 1322
stage monitor systems 1320
stereo cluster loudspeaker system 1261
system documentation 1312
system response curve 1310
systems for religious facilities 1320
talker/listener factors 1251
telescoping shield connection 1304
test equipment 1308
thermal layers of air 1326
Time-Energy-Frequency (TEF) test equipment 1311
transformers 1291, 1305
troubleshooting 1314
unbalanced line 1296
variable Q equalizers 1287
weather-caused deterioration 1327
XLR type connector 1293
Sound system quantities
  units of measurement 1658
Sound Transmission Class (STC) 72
  what's the ear for? 1636
Sound transmission class (STC)
  acoustical modeling and auralization 229
Soundfield FM applications
  interpretation and tour group systems 1541
Sound-field-proportionate factors
  acoustics for auditoriums and concert halls 152
Soundweb
  digital audio interfacing and networks 1486
Source
  acoustical modeling and auralization 218
  fiber optics 483
Source impedance
  audio transformers 285, 287
Source Independent Measurement (SIM) test equipment
  sound system design 1312
Spanning tree
  digital audio interfacing and networks 1496
Spatial filtering
  psychoacoustics 44, 60
Spatial hearing
  psychoacoustics 57
Spatial impression measure R
  acoustics for auditoriums and concert halls 160
SPDIF 1443
Specific flow resistance
  acoustics for auditoriums and concert halls 184
Specific thermal resistance
  heatsinks 372
Specifications for sound reinforcement
  sound system design 1280
Spectator sports
  designing for speech intelligibility 1392
Spectral leakage
  test and measurement 1622
Spectral response
  fundamentals of audio and acoustics 27
Spectrogram
  psychoacoustics 51, 57
Spectrum analysis
  loudspeakers 638
Spectrum analyzers
  VI meters 1002
Speech intelligibility
  designing for speech intelligibility 1387
  signal-to-noise ratio (SNR) 1390
Speech Interference Level
  designing for speech intelligibility 1403
Speech spectrum
  designing for speech intelligibility 1388
Speech Transmission Index (STI)
  acoustics for auditoriums and concert halls 155
  designing for speech intelligibility 1392, 1405
  loudspeaker clusters 654
  test and measurement 1627
Speech Transmission Index—STI
  designing for speech intelligibility 1397
Speech waveforms
  designing for speech intelligibility 1387
Speed of light
  fundamentals of audio and acoustics 27
Speed of sound
  outdoor sound systems 204
Speed/Velocity
  units of measurement 1651
Spherical head model
  psychoacoustics 44, 58

Spherical stylus
  analog discs 1031
Spider
  loudspeakers 598–599
Spiral 425
Spiral shields
  wire and cable 425
SPL
  computer aided sound system design 1376
Splice
  fiber optics 474–475, 477, 483
splice 1071
Splice loss
  fiber optics 460
Split cluster loudspeaker system
  sound system design 1261
Spoken-drama theaters
  acoustics for auditoriums and concert halls 163
Spontaneous emission
  fiber optics 483
Sports halls, gymnasiums
  acoustics for auditoriums and concert halls 164
Sports stadiums and outdoor systems
  sound system design 1322
Spurious emissions
  microphones 567
SRC
  digital audio interfacing and networks 1463
SSM2018
  integrated circuits 343
SSM2141
  integrated circuits 355
ST connector
  fiber optics 483
Stability
  consoles 887
  DSP technology 1164
Stage monitor systems
  Sound system design 1320
Standard MIDI files
  MIDI 1125
Standard STC Contour 73
Standards
  Advanced Encryption Standard (AES) 1153
  AES10 1484
  AES-10–MADI 987
  AES11 1463
  AES2-1984 (r2003) 1345
  AES3 986, 1463, 1481
  AES3id 1473
  AES-42 986
  AES42 1467, 1474, 1481
  AES-4id-2001 118
  AES5 1462
  ANSI Standard S3.5 1332
  ANSI Standard S3.5 1969 1404
  ASTM C423 98
  ASTM E84 122
  AT&T 258A 1500
  DIN 45570 T1 1346
  EIA RS-426A 1277
  EIA standard RS.221.A 521
  EIA standard SE-10 517
  EIA/TIA 568A/568B 1500
  EN 60118 1551
  EN ISO 9921 Feb. 2004 156
  IEC 268-5 (1972) 1346
  IEC 60268-12 1472
  IEC 60958 1482
  IEC 61607-1 1484
  IEC118 1551
  IEEE 1394 1487
  IEEE 1588 1511
  IEEE 802.11 1490
  IEEE-1394 410
  International Building Code 122
  ISO 11654 101
  ISO 15664 118
  ISO 354 98
  ISO/IEC 11801 412
  ITU-R BT.601 423
  Multimedia Home Platform (DVB-MHP) 1154
  National Electrical Code (NEC) 438
  NFPA 286 122
  SMPTE 259M 423
  SMPTE 292M 423
  SMPTE 344M 423
  Standard DIN 45570 1345
  Standard ISO 17497-1 1354
  Standard ISO 354 1354
  TIA/EIA 568A 412
Standars
  SMPTE Standard 428-3-2006 1601
Standing waves
  grounding and interfacing 1185
Standing-wave ratio (SWR)
  grounding and interfacing 1185
Stands and booms
  microphones 581
Stapes
  personal monitors 1433
  psychoacoustics 43
Star coupler
  fiber optics 483
Star network topology
  console 992
State variable filter
  filters and equalizers 795
State-variable filter
  integrated circuits 346
Static Contact Resistance (SCR)
  relays 381
Static fields
  grounding and interfacing 1181
Static line regulation
  power supplies 672
Station intercommunications
  intercoms 1568
Statistical Energy Analysis (SEA)
  acoustical modeling and auralization 228
Status flags
  digital audio interfacing and networks 1478
STC number
  acoustical noise control 72
Steady-state waves
  acoustical modeling and auralization 218
Step index fiber
  fiber optics 483
Step-down auto-transformer
  audio transformers 281
Step-down transformer
  audio transformers 276, 298
  solid state devices 331
Step-up transformer
  analog discs 1027
  audio transformers 276, 298

Stereo boom microphone 548
Stereo cluster loudspeaker system
  sound system design 1261
Stereo consoles 829
Stereo disc groove
  analog discs 1031
Stereo image
  personal monitors 1429
Stereo mic'ing techniques 582
Stereo microphones 542
Stereo mixes
  personal monitors 1422
Stereo monitor mix
  personal monitors 1422
Stereo multiplexed transmission
  personal monitors 1428
Stereo wireless transmission
  personal monitors 1428
Stereo wireless transmitters
  personal monitors 1428
Stevens rule
  psychoacoustics 56
STI
  designing for speech intelligibility 1405
  message repeaters and evacuation systems 1526
  Test and measurement 1627
STI modulation reduction factor
  designing for speech intelligibility 1408
STI/RaSTI
  designing for speech intelligibility 1408
Sticking (contacts)
  relays 377
Stimulated emission
  fiber optics 483
STIPA
  designing for speech intelligibility 1405, 1407
STIPa
  designing for speech intelligibility 1407
Stop band
  filters and equalizers 785
Streaming compression codecs
  consoles 975
Strength measure G
  acoustics for auditoriums and concert halls 153
Stroboscopic disc
  analog discs 1019–1020
Structure coefficient
  acoustics for auditoriums and concert halls 184
Structured cabling
  wire and cable 412
Styli
  analog discs 1031
Stylus
  cantilever
    analog discs 1032
  characteristics
    analog discs 1032
  compliance 1032
  elliptical 1031
  spherical 1031
  stylus tip 1031
  vertical resonance 1033
Stylus tip
  analog discs 1031
Subcode
  optical disc formats 1135
Subframe format
  digital audio interfacing and networks 1464
Subgrouping
  consoles 833
Subjective intelligibility tests
  acoustics for auditoriums and concert halls 156
substrate 1053
Subtractive feedback gate
  consoles 906
Successive-approximation encoder
  consoles 956
Sum-difference networks
  integrated circuits 356
Summation of modal effect
  small room acoustics 134
Summing amplifiers
  attenuators 781
Summing inverter amplifier
  amplifiers 340
Summing Modules 930
Super Audio CD
  optical disc formats 1133
Super Audio Compact Disc (SACD)
  Optical disc formats 1144
Super VHS 421
Supercardioid microphone 495
Superluminescent diodes (SLDs)
  fiber optics 466, 483
Superposition
  fundamentals of audio and acoustics 30
Suppressor grid
  tubes 311
Surface area and volume equations 1678
Surface emitting LED
  fiber optics 463
Surface materials
  acoustical modeling and auralization 219
Surface mount (SMD) relays 386
Surface shapes
  fundamentals of audio and acoustics 29
Surge protection
  grounding and interfacing 1216
Surge voltage
  on a capacitor 259
Surroound sooound
  Pro Logic decoding 1596
Surroound sound
  equal-loudness contour curves 1599
Surround
  loudspeakers 598
Surround panning
  consoles 916
Surround sooound
  audio/video receivers 1598
Surround sound
  5.1 surround sound 1593
  7.1 surround sound format 1598
  Audyssey Dynamic EQ 1600
  channel definitions for digital cinema 1601
  CinemaScope 1593
  Digital Theater Sound (TDS) 1595
  Dolby Digital 1595
  Dolby Digital 5.1 1597
  Dolby Digital Surrond EX 1595
  Dolby Pro logic II 1598

  Dolby Stereo 1593
  LFE 1595
  multichannel magnetic-stripe formats 1593
  multichannel optical format 1593
  optical soundtracks 1593
  Sony Dynamic Digital Sound (SDDS) 1595
  THX Spectral Balancing 1600
  Todd-AO 1593
Surround sound analyzer
  VI meters 1006–1007
Surround sound microphone system 550
Suspension compliance
  microphones 579
Suspension methods
  loudspeakers 599
S-Video
  display technologies 1582
S-video
  wire and cable 421
Swept frequency equalizer
  filters and equalizers 801
Swept sine measurement
  loudspeakers 641
Swinging input control
  consoles 892
Swinging input equalizer
  consoles 893
Swinging output control
  consoles 892
Switch AT
  relays 377
Switch debouncing
  consoles 926
Switched capacitor filter
  filters and equalizers 797
Switches 1495
  digital audio interfacing and networks 1493
Switching current
  relays 377
Switching diodes
  solid state devices 321
Switching regulators
  power supplies 682
Switching voltage
  relays 377
Symbol prefixes
  units of measurement 1647
Symbols
  tubes 311
Symmetrical FIRs
  consoles 965
Symmetrical power
  grounding and interfacing 1215
Symmetrical push-pull transducer microphone 512
Syn-Aud-Con 5
Synchronization
  digital audio interfacing and networks 1474, 1477, 1486
  virtual systems 1443
Synchronous connections
  digital audio interfacing and networks 1461
Synchronous rectification low-voltage power supplies 684
Synchronous rectifiers
  power supplies 688
Synthesizer
  MIDI 1114
System delay
  DSP technology 1174
System distortion
  designing for speech intelligibility 1388
System documentation
  sound system design 1312
System for Improved Acoustic Performance (SIAP)
  acoustics for auditoriums and concert halls 194
System grounding
  grounding and interfacing 1181
System level architecture
  consoles 871, 933
System messages
  MIDI 1108
System noise
  grounding and interfacing 1193
System planning for multi-channel wireless systems
  microphones 571
System response curve
  sound system design 1310
Systems for religious facilities
  sound system design 1320

T

T (or tee) coupler
  fiber optics 483
T attenuators 773
T network bandpass
  filters and equalizers 791
Talkback
  consoles 829
Talker articulation
  designing for speech intelligibility 1401
Talker/listener factors
  sound system design 1251
Tangential error
  analog discs 1022
Tangential mode
  small room acoustics 128
Tangential room modes
  small room acoustics 128
Tangential tracking
  analog discs 1021
Tape delay
  delay 810
Tape guides 1052
Tape metering 1043
Tape recorder transport, maintenance 1088
Tape record-playback equipment equalization standards
  magnetic recording 1056
Tape storage
  magnetic recording 1087
Tape tensioning
  magnetic recording 1047
Tape testing
  tape testingording 1093
Tape transports 1043
Tapped delay
  delay 809
Task of variable acoustics
  acoustics for auditoriums and concert halls 189
T-coil
  assistive listening systems 1546
T-coil hearing aids
  assistive listening systems 1551
TCP/IP platform 989
  consoles 989

TDS
  test and measurement 1616
TEF 20 67
TEF 25 70
TEF25™ 1626
Teflon
  wire and cable 406
Telcom Closet (TC)
  wire and cable 412
Teleconferencing
  preamplifiers and mixers 754
  sound system design 1329
Teleconferencing equipment
  preamplifiers and mixers 758
Telephone hybrid circuit
  audio transformers 300
Telephone interface
  intercoms 1575
  preamplifiers and mixers 758
Telephone isolation
  audio transformers 300
Telephone/transmission system interface
  preamplifiers and mixers 757
Telescopic grounds
  wire and cable 428
Telescoping shield connection
  sound system design 1304
Temperature (T)
  units of measurement 1648, 1652
Temperature coefficient of resistance (TCR) 243
Temperature differential
  heatsinks 369
Temperature gradient
  heatsinks 367
Temporal bone
  psychoacoustics 44
Temporal coding
  psychoacoustics 55–56
Temporary hearing damage
  what's the ear for? 1635
Temporary Threshold Shift (TTS)
  personal monitors 1434
Temporary threshold shift (TTS)
  what's the ear for? 1633, 1635
Terminating connectors
  wire and cable 416
Termination impedance
  consoles 859
Test and Measurement
  electrical testing 1607
Test and measurement
  %Alcons 1625
  acoustic wavelengths 1619
  acoustical testing 1607
  amplitude versus time 1611
  anechoic measurements 1614
  Blackman-Harris 1623
  critical distance 1619
  data windows 1622
  digital sampling 1620
  domain chart 1613
  dual-channel FFT analyzers 1616
  early-decay time (EDT) 1627
  EDT 1627
  electrical test instruments 1607
  Envelope-Time Curve 1618
  Envelope-Time Curve (ETC) 1612, 1616
  ETC 1616
  Fast Fourier Transform 1615
  FFT measurements 1615
  Fourier transform 1613
  frequency response 1613
  gated-noise test 1611
  Hamming 1623
  Hann window 1623
  Hilbert Transform 1618
  human hearing system 1619
  human perception 1625
  IFFT 1615
  impulse response 1612–1613, 1618
  impulse response of a loudspeaker/room combination 1620
  impulsed response tests 1628
  Inverse Fast Fourier Transform 1615
  maximum-length sequence (MLS) 1616
  measurement system 1613
  microphone calibrator 1610
  microphone placement 1619
  MLS 1616
  noise criteria ratings 1610
  phase response 1613
  polarity 1627
  polarity testers 1627
  reflection-free measurements 1614
  reverberant field 1620
  reverberation time 1611
  short interval reflections 1609
  SLM 1609
  sound level measurements 1608–1609
  sound level meter 1608
  sound pressure level 1608
  sound pressure measurements 1608
  spectral leakage 1622
  Speech Transmission Index (STI) 1627
  STI 1627
  TDS 1616
  time domain response 1618
  time period 1613
  time window 1621
  time window length 1618
  Time-Delay Spectrometry (TDS) 1616
  transfer function 1612–1613
  transforms 1613
Test equipment
  sound system design 1308
Testing
  audio transformers 302
Tetrode
  tubes 311
TFT-LCD technology
  display technologies 1584
THAT 2180
  integrated circuits 343
THAT1200-series
  integrated circuits 355
THAT1510
  integrated circuits 351
THAT2252
  integrated circuits 347
THD 284, 1472
  audio transformers 284
  magnetic recording 1059
Thermal characteristics for the heat sink 371
Thermal conductivity
  heatsinks 370

Thermal failure
  loudspeakers 609
Thermal layers of air
  sound system design 1326
Thermal limit
  loudspeakers 607
Thermal noise
  fiber optics 468
  gain structure 1225–1226
  microphones 518
  preamplifiers and mixers 737
  units of measurement 1661
Thermal performance
  heatsinks 367
Thermal regulation
  power supplies 672
Thermal relaxation
  acoustical modeling and auralization 217
Thermal resistance
  heatsinks 370–372
  loudspeakers 608–609
Thermal solutions
  heatsinks 368–369
Thermal spreading resistance
  heatsinks 368
Thermally conductive adhesives
  heatsinks 371
Thermistors (thermal-sensitive resistors) 249
Thermocouples 249
Thermoset compounds
  wire and cable 406
Thevenin's theorem
  units of measurement 1653
Thiele-Small parameters
  loudspeakers 637
Thomas B. Doolittle 8
Thomas Stockham 8
Three chip DLP projection TV
  display technologies 1586
Three entry cardioid microphone 500
Three-beam design pickup
  optical disc formats 1138
Three-phase power supplies 674
Threshold current
  fiber optics 483
Threshold of hearing
  psychoacoustics 50
Thresholding
  consoles 908
Thresholds
  psychoacoustics 50
THX Spectral Balancing
  surround sound 1600
Thyristors
  solid state devices 322
TIA and ISO classifications for structured cabling
  wire and cable 412
Tight buffer fiber jackets
  fiber optics 457
Timbre
  psychoacoustics 57
Time (t)
  units of measurement 1648
Time and distance considerations 217
Time code
  MIDI 1128
Time constant
  capacitors 253
  inductors 267
Time criteria
  acoustics for auditoriums and concert halls 148
Time Delay Spectrometry
  loudspeakers 641
Time domain
  DSP technology 1165
  preamplifiers and mixers 738
Time Domain Reflectometer
  wire and cable 432
Time domain response
  test and measurement 1618
Time invariance
  DSP technology 1164
Time manipulation
  consoles 961
Time period
  fundamentals of audio and acoustics 27
  test and measurement 1613
Time processing in DSP
  consoles 961
Time response
  fundamentals of audio and acoustics 34
Time window
  psychoacoustics 51
  test and measurement 1621
Time window length
  test and measurement 1618
Time-arrivals
  computer aided sound system design 1378
Timebase corrector
  magnetic recording 1080
Timecode
  magnetic recording 1087
Time-Delay Spectrometry (TDS)
  test and measurement 1616
Time-domain effects
  consoles 839
Time-Energy-Frequency (TEF) test equipment
  sound system design 1311
Time-windowed measurements
  loudspeakers 639
Tinnitus
  fundamentals of audio and acoustics 35
  personal monitors 1434
  psychoacoustics 50
  what's the ear for? 1635
Todd-AO
  surround sound 1593
TOLED
  display technologies 1588
Tom Osborne 18
Tonal coloration
  fundamentals of audio and acoustics 31
Tonearms
  analog discs 1021
  antiskating force 1025
  effective tonearm length 1023
  geometry 1021
  linear speed 1023
  resonant damping 1025
  skating force 1023
  styli 1026
  tangential error 1022
  tangential tracking 1021
  vertical tracking angle 1025
Tools of simulation
  computer aided sound system design 1370
Toroidal core

  audio transformers 280
Toroidal power transformers
  Audio transformers 283
Total absorption
  units of measurement 1661
Total harmonic distortion
  audio transformers 284
Total internal reflection
  fiber optics 483
Totalizer/preset counter
  relays 391
Totem pole
  consoles 880
Tour group system
  interpretation and tour group systems 1538
Tracking correction signal
  optical disc formats 1139
Transceivers
  fiber optics 469
Transconductance
  tubes 312
Transconductance amplifier 703
Transducer types
  loudspeakers 598
Transducers
  balanced armature 1424
  dual driver earphones 1424
  dynamic earphones 1424
Transfer function
  amplifier 704
  loudspeakers 629
  test and measurement 1612–1613
Transfer functions
  loudspeakers 624
Transformer
  step-up 1027
  step-up transformer 331
Transformer characteristics
  consoles 862
Transformer input microphone preamplifiers
  integrated circuits 349
Transformer-coupled solid-state relays 390
Transformerless front ends
  consoles 866
Transformers 1063
  basics 275
  consoles 867
  preamplifiers and mixers 739
  see Audio transformers 277
  sound system design 1291, 1305
  windings 276
Transforms
  test and measurement 1613
Transient intermodulation distortion (TID)
  consoles 840
Transient response
  filters and equalizers 786
  grounding and interfacing 1212
Transistor
  bipolar junction 343
  MOSFET 720
Transistor mounting kit
  heatsinks 371
Transistors
  alloy-junction 325
  base-current curves 333
  bias circuits 329
  breakdown voltage 332
  complementary symmetry monolithic Darlington power transistors 719
  consoles 834
  current flow 328
  double-diffused epitaxial mesa 326
  drift-field 325
  equivalent circuits 328
  field effect 326
  forward-current-transfer ratio 329
  gain-bandwidth product 329
  grown-junction 325
  input resistance 328
  insulated-gate 327
  load line 333–334
  mesa 325
  microalloy diffused 325
  MOSFET 327, 335
  noise figure 331
  planar 326
  polarity 328
  punch through 332
  small and large-signal characteristics 330
  solid state devices 325
Transition band
  filters and equalizers 785
Transmission characteristics
  audio transformers 302
Transmission coefficient
  acoustics for auditoriums and concert halls 183
Transmission line
  relays 377
Transmission line termination
  wire and cable 435
Transmission lines
  grounding and interfacing 1184
Transmission loss
  acoustical noise control 71
  fiber optics 483
Transmission loss of a compound barrier
  acoustical noise control 87
Transmission loss of single glass plates
  acoustical noise control 84
Transmission loss of spaced glass plates
  acoustical noise control 85
Transmitter
  fiber optics 483
  specifications
    fiber optics 476
Transmitter antenna
  microphones 567
Transmitter/receiver modules
  fiber optics 469
Transparent OLED
  display technologies 1588
Transportation services applications
  message repeaters and evacuation systems 1518
Transresistance amplifier 703
Transversal equalizers
  filters and equalizers 801
Transversal filter
  optical disc formats 1140
Transversal filters
  consoles 963
TRAP horns 650
  loudspeaker clusters 650
Trapezoidal ground plan
  acoustics for auditoriums and

- concert halls 169
- Travel-Time-Based Reverberation-Time Enhancing Systems
  - acoustics for auditoriums and concert halls 191
- Treble horns
  - computer aided sound system design 1350
- Tree coupler
  - fiber optics 483
- Triacs
  - solid state devices 325
- Triaxial cable
  - wire and cable 421
- Triboelectric noise
  - wire and cable 430
- Triode
  - tubes 311, 313
- Tri-polar array
  - Loudspeaker clusters 653
- Troubleshooting a sound system 1314
- True Array Principle (TRAP)
  - loudspeaker clusters 651
- TTL (Transistor-Transistor Logic) 746
- Tubes 311
  - amplification factor 312
  - base diagrams 311
  - beam- power 311
  - cathode 311
  - control grid 311, 314
  - diode 311
  - filament 311
  - Fleming valve 311
  - grid bias 314
  - grid voltage 315
  - grid-current 315
  - grid-to-plate capacitance 314
  - heater 317
  - heptode 311
  - hexode 311
  - input capacitance 313
  - internal capacitance 313
  - nomenclature 312
  - output capacitance 314
  - pentagrid 311
  - pentode 311
  - plate 311
  - plate dissipation 315
  - plate efficiency 315
  - plate resistance 314
  - polarity 313
  - power sensitivity 315
  - screen grid 311, 315
  - step-up transformer 331
  - suppressor grid 311
  - symbols 311
  - tetrode 311
  - transconductance 312
  - triode 311, 313
  - tube elements 311
  - types 311
  - voltage gain 312
- Tuned dipolar array
  - loudspeaker clusters 653
- Tuned stub noise attenuators
  - acoustical noise control 92
- Tunnel diode
  - solid state devices 322
- Turnoff time
  - solid state devices 324
- Turnon time
  - solid state devices 324
- Turns ratio
  - audio transformers 276–277, 280, 299, 302
- Turntable
  - acoustic noise 1020
  - analog discs 1018
  - cartridge voltage output 1033
  - cleaning records 1037
  - drive systems 1018
  - elliptical stylus 1031
  - induced-magnet cartridges 1028
  - laser 1034
  - magnetic field 1027–1028
  - moving-coil cartridges 1027
  - moving-magnet cartridges 1028
  - phono preamplifiers 1033
  - platter 1018
  - semiconductor pickup cartridges 1029
  - speed of rotation 1019
  - spherical stylus 1031
  - step-up transformer 1027–1028
  - styli 1026, 1031
  - stylus cantilever 1032
  - stylus characteristics 1032
  - stylus tip 1031
  - tonearm resonance damping 1025
  - tonearms 1021
  - variable-reluctance cartridges 1027
- Twisted pair
  - wire and cable 407
- Twisted pair impedance
  - wire and cable 434
- Two channel party line system
  - intercoms 1570
- Two-integrator loop
  - consoles 885
- Two-way cardioid microphone 502
- Two-way crossovers
  - loudspeakers 624
- Two-wire telephone connection
  - preamplifiers and mixers 758
- txUnicastMode
  - digital audio interfacing and networks 1506
- Tympanic membrane
  - psychoacoustics 43
- Types of automatic microphone mixers
  - preamplifiers and mixers 748
- Types of tubes 311
- Typical digital circuits 363
- Typical sound pressure levels
  - fundamentals of audio and acoustics 26

U

- U attenuators 775
- U.S. to SI units conversion factors
  - units of measurement 1666
- U50
  - designing for speech intelligibility 1405
- U80
  - designing for speech intelligibility 1405
- UDF
  - optical disc formats 1146
- UFC
  - message repeaters and evacuation systems 1526

Ultra capacitor characteristics
  power supplies 691
Ultra capacitor UPS
  power supplies 690
Ultrasound
  message repeaters and evacuation systems 1524
Unbalanced
  microphones 521
Unbalanced cable
  grounding and interfacing 1202
Unbalanced electrical interface
  digital audio interfacing and networks 1483
Unbalanced input
  audio transformers 297–298
Unbalanced interface
  audio transformers 296
  grounding and interfacing 1213
Unbalanced line
  sound system design 1296
Unbalanced output
  audio transformers 297–298
Unbalanced pad 767
Under water loudspeakers 605
Underhung coil
  loudspeakers 607
UniCam Connector System tool
  fiber optics 474
Unicast bundles
  digital audio interfacing and networks 1506
Unidirectional microphone 493, 495, 498, 506
Units of Measurement 1647
Units of measurement 1647
  %Alcons 1659
  abbreviations 1667
  absorption coefficient 1660
  acceleration. 1652
  amplitude 1652
  audible frequency range 1679
  audio frequency range 1666
  audio spectrum 1678
  audio standards 1663
  average absorption coefficient 1660
  capacitance (C) 1650
  conductance 1650
  critical distance 1658
  critical distance modifier (M) 1659
  current (I) 1648
  decibel (dB) 1655
  derived quantities in physics 1648
  directivity factor (Q) 1660
  displacement 1651
  electric capacitance (C) 1650
  electric charge (Q) 1650
  electric conductance 1650
  electric inductance 1651
  electric potential difference (V) 1650
  electric resistance (R) 1650
  electrical power required (EPR) 1660
  energy (E) 1649
  equivalent acoustic distance (EAD) 1658
  equivalent input noise (EIN) 1661
  feedback stability margin (FSM) 1659
  force (F) 1650
  frequency (f) 1651
  fundamental quantities in physics 1647
  Greek alphabet 1663
  inductance 1651
  ISO numbers 1662
  kinetic energy (T) 1649
  Kirchhoff's current law 1654
  Kirchhoff's laws 1653
  Kirchhoff's voltage law (KVL) 1653
  length (l) 1648
  limiting distance 1658
  loudspeaker sensitivity 1661
  luminous flux 1651
  luminous flux density 1651
  luminous intensity (IL) 1649
  magnetic flux 1651
  magnetic flux density 1651
  mass (M) 1648
  max program level 1661
  mean-free path (MFP) 1660
  molecular substance (n) 1649
  multiple and submultiple prefixes 1647
  needed acoustic gain (NAG) 1659
  number of open microphones (NOM) 1659
  Ohm's law 1654
  phase 1652
  phase angle 1649, 1652
  phase difference 1652
  phase shift 1652
  phase velocity 1652
  potential acoustic gain (PAG) 1659
  potential energy (U) 1649
  power (W) 1650
  pressure 1650
  radio frequency spectrum 1654
  resistance 1650
  SI system 1647
  signal delay 1662
  signal to noise ratio (SNR) 1660
  solid angle (A) 1649
  sound intensity 1651
  sound pressure level (SPL) 1657, 1660
  sound system quantities 1658
  speed/velocity 1651
  surface area and volume equations 1678
  symbol prefixes 1647
  temperature (T) 1648, 1652
  thermal noise 1661
  Thevenin's theorem 1653
  time (t) 1648
  total absorption 1661
  U.S. to SI units conversion factors 1666
  wavelength (M) 1652
  weight 1652
Unity
  gain structure 1225
Unity amplification method
  gain structure 1228
Unity gain Sallen-Key filter
  filters and equalizers 795
Unity-gain current-summing amplifiers 781
Universal Disc Format (UDF) Bridge
  optical disc formats 1146

Universal Disk Format (CD-UDF)
  Optical disc formats 1143
Universal mix buses
  consoles 978
Universal serial bus (USB)
  wire and cable 411
Universal winding
  audio transformers 282
Unrelated crosstalk
  consoles 919
Unshielded output transformer
  audio transformers 283
Unshielded twisted pair (UTP)
  digital audio interfacing and networks 1470
USB
  consoles 988
USB microphones 556
  consoles 989
User Defined Protocol (UDP)
  consoles 991

V

Vacuum tubes 311
Varactor diode
  solid state devices 322
Variable acoustics
  acoustics for auditoriums and concert halls 188
Variable Q
  loudspeaker clusters 662
Variable Q equalizers
  sound system design 1287
Variable reluctance cartridges
  analog discs 1027
Variable threshold automatic microphone mixers 748
Vario-acoustics
  acoustics for auditoriums and concert halls 188
Varistors (voltage-sensitive resistors) 249
VCA
  integrated circuits 341
VCA channel application
  consoles 939
VCSEL
  fiber optics 466
Velocity of propagation
  wire and cable 421, 424
Vented boxes
  loudspeakers 636
Vertical cavity surface emitting laser (VCSEL)
  fiber optics 466, 483
Vertical resonance
  analog discs 1033
  stylus 1033
Vertical tracking angle
  analog discs 1025
Vertical woofer arrays
  loudspeaker clusters 653
VGA
  Wire and cable 422
VI meter
  fundamentals of audio and acoustics 36
VI meters
  +4 dBu 997
  -10 dBV 997
  AES/EBU digital peak meter 1004
  analog peak program meter 1004
  attenuators 1000
  average power 1006
  Bit Stream Corruption 1006
  British Broadcast Standard 1004
  dB FS 997
  dBA 997
  dBC 997
  dBf 997
  dBm 997
  dBr 997
  dB-SPL 997
  dBu 997
  dBV 997
  dbW 997
  digital peak program 1004
  DIN 45406 Standard 1003
  DIN Scale 997
  EBU standard 1005
  harmonic distortion 998
  loudness 1004–1005
  loudness meter 1004
  Nordic N9 Standard 1004
  peak program 1003
  persistence 1006
  phase correlation meter 1004
  phase error 1006
  power-level 1003
  power-output 1003
  relative loudness 1006
  SMPTE 1005
  spectrum analyzers 1002
  surround sound analyzer 1006–1007
Vibration isolation
  acoustical noise control 90
Video codecs
  optical disc formats 1152
Video Graphics Array (VGA)
  display technologies 1582
Video triaxial cable
  Wire and cable 421
Vinyl
  wire and cable 405
Virtual Local Area Network (VLAN)
  digital audio interfacing and networks 1496
Virtual pitch
  psychoacoustics 57
Virtual reality
  acoustical modeling and auralization 233
Virtual sound processors
  virtual systems 1446
Virtual sound systems
  virtual systems 1452
Virtual Studio Technology
  consoles 984
Virtual systems
  active acoustics 1455
  CDRW 1439
  CobraNet 1443, 1447
  compressors 1439, 1445
  consoles 1440, 1445, 1452
  crossover 1449
  delay 1451
  digital audio processors 1445
  digitally controlled analog devices 1440
  digitally controlled digital audio devices 1441
  distortion 1442–1443
  DSP 1440, 1443, 1447, 1451–1453
  DVD 1439
  dynamic processors 1439

  dynamic range 1442
  equalization 1439
  ethernet 1448
  FIR filters 1454
  level control 1439
  level matching 1442
  loudspeakers 1453
  MediaMatrix 1447
  microphone 1452
  multifunction devices 1444
  processing system 1454
  Sample Rate Converter 1443
  synchronization 1443
  virtual sound processors 1446
  virtual sound systems 1452
Virtual-earth mixers
  consoles 910
Virtual-earth mixing buses
  consoles 932
Viscoelastic
  loudspeakers 599
VLSI
  DSP technology 1173
Voice coil
  loudspeakers 598
Voice evacuation
  message repeaters and evacuation systems 1526
Voice evacuation/mass notification systems
  message repeaters and evacuation systems 1526
Voice over IP
  intercoms 1572
Voice warning horns
  loudspeakers 613
VOIP
  intercoms 1566
Volt
  fundamentals of audio and acoustics 32
Voltage amplifier 703
Voltage amplitude
  gain structure 1227
Voltage coefficient 243
Voltage controlled amplifier (VCA)
  consoles 936
Voltage controlled amplifiers
  integrated circuits 342
Voltage controlled crystal oscillator
  digital audio interfacing and networks 1474
Voltage divider
  grounding and interfacing 1186
Voltage divider capacitor microphone 508
Voltage gain
  tubes 312
Voltage matching
  grounding and interfacing 1186
Voltage ratio
  audio transformers 276
Voltage standing wave ratio (VSWR)
  relays 377, 385
Voltage-controlled amplifier (VCA) 779
  consoles 936
Voltage-controlled-attenuator
  consoles 900
Volume control 776
Volume indicator (VI) meter 997
VU
  Audio and Acoustic DNA 10
VU meter 933
  consoles 920
  gain structure 1228
VU meters 997
  bar graph 1002
  characteristic impedance 997
  crest factor 999
  D'Arsonval meter 1000
  D'Arsonval movement 998
  dynamic characteristics 998
  impedance 998, 1000
  Impedance Level Correction 1000
  instrument scale 998
  line impedance 1001
  logarithmic amplifier 1002
  meter ballistics 999
  moving coil VU meter 998
  NTSC 1002
  reference levels 1000
  sensitivity 998
  UREI) Model 970 Vidigraf 1002
  voltages at various impedances 1001
  volume unit 997
  Volume-Unit meter (VU) 998
  wide-range 1002

W

W.C. Sabine
  acoustical treatment for indoor areas 97
Wake-up time
  consoles 905
Wall caulking
  acoustical noise control 77
Wall construction
  acoustical noise control 74
Wall materials
  computer aided sound system design 1354
Wallace Clement Sabine 14
waterproof wireless microphone systems 563
WAV files
  optical disc formats 1135
Wave Acoustics routines
  computer aided sound system design 1354
Wave equation
  acoustical modeling and auralization 225
Wave equation models
  acoustical modeling and auralization 225
Wavelength
  fiber optics 452, 464, 466
  fundamentals of audio and acoustics 27
  peak
    fiber optics 482
Wavelength (M)
  units of measurement 1652
Wavelength division multiplexing (WDM)
  fiber optics 484
Wavelet analysis
  psychoacoustics 51
Wayne Rudmose 13
Weather-caused deterioration
  sound system design 1327
Weber fraction in dB
  psychoacoustics 54
Weber's Law
  psychoacoustics 54
Weight

  units of measurement 1652
Weighting curves
  fundamentals of audio and acoustics 34
Weighting scales
  fundamentals of audio and acoustics 35
Wein Bridge
  consoles 885
Western Electric 11
Western Electric Sound Products Division 12
What's the ear for?
  analog A and C weighted filters 1639
  analog Z weighting (linear) filter 1639
  annoyance factor 1637
  community noise problems 1636
  dBC 1639
  earmuffs 1635
  earplugs 1635
  EPA 1633
  hearing damage 1633, 1635
  hearing loss 1633
  hearing tests 1635
  Leq 1640
  NIOSH 1633
  noise 1636
  OSHA 1633
  OSHA permitted noise levels 1634
  permanent threshold shift (PTS) 1633
  protecting concert-goers 1636
  protecting hearing 1635
  protecting the audience 1642
  protecting the community 1636
  SLARM™ 1638
  SLARMSolution™ 1637
  SLARMSolution™ applications 1641
  sound transmission class (STC) 1636
  temporary hearing damage 1635
  temporary threshold shift (TTS) 1633, 1635
  tinnitus 1635
  WHO 1633
Wheeler's equation
  assistive listening systems 1549
White noise
  psychoacoustics 57
WHO
  what's the ear for? 1633
William Thomson 5
Wind noise reduction figures
  microphones 576
Winding resistance
  audio transformers 280
  consoles 883
Windings
  audio transformers 276
Window
  fiber optics 484
Windowing
  consoles 964
Windows
  acoustical noise control 84
Windows Media Audio (WMA)
  digital audio interfacing and networks 1461
Windscreens
  microphones 576
Wire
  consoles 845
  grounding and interfacing 1183
Wire and cable
  70.7 volt loudspeaker system 435
  ac current in various countries 445
  ac power cords and receptacles 443
  adapters 417
  AES/EBU digital audio cable 430
  AES/EBU digital coaxial cable 430
  alien crosstalk (ANEXT) 416
  attenuation 432
  backbone cabling 412
  balanced lines 407
  baluns 416
  braid shields 426
  Canadian Electrical Code cable substitution chart 441
  Category 1 413
  Category 2 413
  Category 3 413
  Category 4 413
  Category 5 413
  Category 5e 413
  Category 6 413
  Category 6 F/UTP 413
  Category 6a 413
  Category 7 S/STP 413
  category cables 412
  CATV broadcast cable 419
  CCTV (closed circuit television) 419
  characteristic impedance 432–433, 437
  coaxial 418
  coaxial cable construction 418
  coaxial cable installation 419
  coaxial cable signal loss 433
  coaxial cable termination 420
  color code for nonpaired cables 407
  color code for UTP cable 415
  color codes for paired cables 408
  color codes for snake cables 409
  combination shields 426
  conduit fill 431
  crosstalk 437
  current ratings 403
  damping factor 437
  definitions 412
  delay 431
  delay skew 416, 431
  diameter and bend radius for 10GbE cabling 415
  dielectric constants 403, 425
  digital video 422
  Digital Visual Interface (DVI) 409
  DisplayPort 411
  equal level far-end crosstalk (EL-FEXT) 415
  Ethernet 412
  far-end crosstalk (FEXT) 415
  FireWire serial digital 410
  foil-braid-foil-braid 427
  foils shields 426

French braid shields 426
ground fault interrupters 442
ground loops 428
grounds of shields 427
HD-SDI 421
High Definition Multimedia Interface 410
high-frequency 431
high-frequency, high-data rate 409
horizontal cabling 412
hub 412
insulation 403
insulation characteristics 405
insulation color codes 407
Intermediate Cross-Connect (IXC) 412
international power 445
jackets 405
LAN 416
loading coil 431
loudspeaker 435
Main Cross-Connect (MXC) 412
maximum continuous current rating 402
maximum distances between areas for UPT cable 415
multiconductor 407
multipair cables 408
National Electric Code 438
near-end crosstalk (NEXT) 414
NEC cable groups 440
NEC substitution chart 440
neoprene 406
physiological effects of shock current on humans 443
plastics 405
plenum cable 441
polyethylene 405
polypropylene 406
POTS 414
Power Over Ethernet (POE) 417
Power Over Ethernet Plus (POE Plus) 417
power-sum near-end crosstalk (PS-NEXT) 415
properties of plastic insulation 404
properties of rubber insulation 404
radio guide designations (RG) 424
receiver quality 423
resistance 397
return loss 420
return loss (RL) 415
RGB cable 421
rubber 406
serial digital video (SDI) 423
shielded power cable 444
shielding 425
shields 426
silicone 406
single conductor 406
skin depth 402
skin effect 402
SMPTE serial digital performance specifications 424
snakes 408
spiral shields 425
structured cabling 412
S-video 421
Teflon 406
Telcom Closet (TC) 412
telescopic grounds 428
terminating connectors 416
thermoset compounds 406
TIA and ISO classifications for structured cabling 412
Time Delay Reflectometer 432
transmission line termination 435
triaxial cable 421
triboelectric noise 430
tri-shield cable 427
twisted pair 407
twisted pair impedance 434
universal serial bus (USB) 411
velocity of propagation 421, 424
VGA 422
video triaxial 421
vinyl 405
Z-Fold foil type shielded wire 426

Wire data
  assistive listening systems 1550
Wireless communication systems
  microphones 557
Wireless frequency bands of operation
  microphones 558
Wireless intercom systems
  intercoms 1574
Wireless intercoms 1569
Wireless IR system receiver
  assistive listening systems 1557
Wireless microphone receiver sensitivity 560
Wireless microphone receiving antenna systems 561
Wireless microphone transmitter considerations 566
Wireless MIDI
  MIDI 1111
Wireless monitor systems
  microphones 571
Wireless receiver
  microphones 568
Wolfgang Ahnert 19

X

XLD connector
  digital audio interfacing and networks 1481
XLR Type Connector
  sound system design 1293
XY stereo
  microphones 543

Y

Y coupler
  fiber optics 484
YPbPr
  display technologies 1582

Z

Zener diode 321
Zero impedance
  consoles 834
Zero impedance amplifier
  audio transformers 288
Zero rejection
  audio transformers 296
Zero-crossing
  consoles 941
Z-Fold foil type shielded wire 426
Zone paging
  intercoms 1564

Zoom microphone 538
Z-transform
DSP technology 1166
μ
audio transformers 278